vsm

VSMDEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

ARTHRITIS ADVISORY COMMITTEE

MEETING

GUIDANCE FOR INDUSTRY: CLINICAL DEVELOPMENT PROGRAMS FOR DRUGS, DEVICES AND BIOLOGICAL PRODUCTS INTENDED FOR THE TREATMENT OF OSTEOARTHRITIS

Friday, February 20, 1998

8:10 a.m.

MILLER REPORTING COMPANY, INC.
507 C Street, N.E.
Washington, D.C. 20002
(202) 546-6666

Bethesda Holiday Inn
8120 Wisconsin Avenue
Bethesda, Maryland

P A R T I C I P A N T S

Members:

Michelle Petri, M.D., M.P.H., Chair
Steven B. Abramson, M.D.
E. Nigel Harris, M.D.
Matthew H. Liang, M.D., M.P.H.
Harvinder S. Luthra, M.D.
Frank Pucino, Jr., Pharm.D.
Barbara C. Tilley, Ph.D.

Consultants:

Leigh F. Callahan, Ph.D.
Marlene Egger, Ph.D.
Felix Fernandez-Madrid, M.D., Ph.D.
Larry W. Moreland, M.D.
Barbara White, M.D.

FDA Staff:

Kathleen Reedy, Executive Secretary

Michael Weintraub, M.D.
Kent R. Johnson, M.D.
Jeffrey N. Siegel, M.D.
William D. Schwieterman, M.D.
James P. Witter, M.D., Ph.D.

C O N T E N T S

P R O C E E D I N G S

CHAIRPERSON PETRI: Good morning. My name is Michelle Petri. I want to welcome everyone to the Arthritis Advisory Committee meeting. Today we are going to be discussing osteoarthritis guidelines. I would like to start by asking the members of the committee to introduce themselves and if we can start at my right and go around.

MS. EGGER: I'm Marlene Egger. I see I'm an honorary M.D. today. I'm a statistician in Salt Lake City.

DR. FERNANDEZ-MADRID: Felix Fernandez-Madrid, Wayne State University.

DR. CALLAHAN: Leigh Callahan, the University of North Carolina in Chapel Hill.

DR. MORELAND: Larry Moreland, the University of Alabama at Birmingham.

DR. WHITE: Barbara White, University of Maryland

DR. HARRIS: I'm Nigel Harris. I'm Dean at Morehouse School of Medicine.

DR. TILLEY: Barbara Tilley, biostatistician from Henry Ford Health System in Detroit.

DR. LUTHRA: I'm Harvy Luthra from Mayo Clinic, Rochester, Minnesota.

DR. ABRAMSON: Steve Abramson from the Hospital for Joint Diseases, NYU.

MS. REEDY: Kathleen Reedy, Food and Drug Administration.

DR. LIANG: I'm Matt Liang from Harvard Medical School.

DR. PUCINO: Frank Pucino from the National Institutes of Health.

DR. WITTER: I'm Jim Witter, FDA.

DR. JOHNSON: Kent Johnson, FDA.

DR. WEINTRAUB: Mike Weintraub, FDA.

CHAIRPERSON PETRI: And I would like to ask Kathleen Reedy, our executive secretary, for the meeting statement.

MS. REEDY: Conflict of interest statement for the Arthritis Advisory Committee, February 20, 1998. The following announcement addresses the issue of conflict of interest with regard to this meeting and is made a part of the record to preclude even the appearance of such at this meeting. Since the issues to be discussed by the committee will not have a unique impact on any particular firm or product but rather may have widespread implications with respect to entire classes of products, in accordance with 18 United States Code 208, waivers have been granted to each member and consultant participating in the committee meeting. A copy of these waiver statements may be obtained

from the agency's Freedom of Information Office, Room 12A30, Parklawn Building.

In the event that the discussions involve any other products or firms not already on the agenda for which an FDA participant has a financial interest, the participants are aware of the need to exclude themselves from such involvement and their exclusion will be noted for the record. With respect to all other participants, we ask in the interest of fairness that they address any current or previous financial involvement with any firm whose products they may wish to comment upon.

CHAIRPERSON PETRI: Thank you, Kathleen. Now I would like to ask Michael Weintraub, Acting Director, Division of Anti-Inflammatory, Analgesic and Ophthalmic Drug Products for his statement.

DR. WEINTRAUB: Good morning. Today we are going to be discussing the guidance for industry on osteoarthritis. The guidance focuses on the development programs for devices, biological products and drugs aimed at treating osteoarthritis. My job this morning is to set the stage for our discussion and to remind you that it is also relatively early in the development of this FDA guidance document.

However, we felt that it would be useful to start

out discussing it and presenting some of our early thoughts in the areas of study, the resulting claims and indications, the clinical trial designs that we could use to arrive at those claims and indications, the risk-benefit relationship, and assessments which are really so important for all three centers because this is an FDA wide document. Its development involves the Center for Devices and Radiologic Health, the Center for Biologics, and of course CDER as well, the Center for Drug Evaluation and Research.

We would like to use the document and the questions as a starting point. But in the discussion, I hope you won't be limited--you shouldn't feel limited or restrained in any way. We want a wide-ranging discussion. Now, in my looking at the document, not having been involved in its creation, so to speak, but I looked at it as a member of the committee, and I think the first thing this document does is lay out some areas of study which really are being enlarged. Of course, there is always pain, but we are enlarging--and we are talking about traditional measures of pain--but we are also discussing enlarging the measurement of pain to the nonsignal joint and thinking about how we can measure those effects, drug effects or device effects or biological product effects, by a patient global score.

There are functional measures as well such as

walking without pain and other measures. And I feel that that is an enlargement of what has been done traditionally. One of the most interesting newer areas of study is the structure of joints. We are very interested in getting your thoughts on looking at the structure of joints. Now there are a variety of ways we can do that and measure the outcome. For example, we can look at cartilage turnover or bone turnover on metabolic basis, measured by direct vision even, perhaps with arthroscopy, or laboratory measures, biochemical markers of bone and cartilage turnover.

Now some of these measures clearly move beyond the traditional patient-derived outcome. However, this document doesn't neglect any important measures of effectiveness, but as I said before, we are expanding our look at measures of effectiveness as well. We're asking you to think about analyses that are more easily understood by physicians and by patients. I believe we should be looking at individual patient outcomes and the use of clinically meaningful hurdles whenever it's logical and feasible. Now, I'm not trying to end the FDA's well known total employment act for statisticians.

[Laughter.]

DR. WEINTRAUB: But I do want to get away from some of the arcane measures and stick to the things that are

easily understood such as clinically meaningful hurdles. Now a number of years ago--as I was looking at this document--a number of years ago I wrote a paper called "The Law of Confounding by Sensible Behavior." Now this wasn't in a major medical journal, but it was a thought piece that I had and had a chance to have published in a journal. One of the things that I discussed is, of course, the communication between patient and physician about a new treatment for--in this case the first one discussed was osteoarthritis--and I said that, look, patients can change their lives in a way which sort of hides or makes more difficult the measurement of drug effects, and I think that that's right, and we have to understand that that's how we bring function into this calculus, into figuring out understanding how to measure these things.

Now, the copies of this paper are available at about $19.95 apiece, but none of you have to--no--

[Laughter.]

DR. WEINTRAUB: The thing is it attempts to integrate the things that I found when I did studies about how patients would say, doc, this drug for my arthritis just doesn't seem to be helping very much, and he would say, but how come you haven't--it seems you haven't been hiring a workman to help you farm? He says, yeah, I'm riding my

tractor, you know, a lot. I haven't had to hire a workman, but my arthritis is just as bad. The patient had difficulty in integrating those two features, that his pain was better so he drove his tractor more and made his pain worse. So it's a difficult problem faced all the time in these kinds of studies.

Now, in discussing the claims and indications, it is another area we have enlarged, and we have enlarged our understanding to include time. We have outlined durations of studies to help establish the indication or to investigate how long an effect may last. We have also looked at changes in the natural history of the disease. For example, the prevention of development of osteoarthritis in previously disease-free joints. We are also looking for changes in the natural history of therapy. For example, what are we doing with joint surgery? Are we able to push that ahead into the future? All of these things, as I see them, are very early thoughts in our development of this guideline. Now we don't expect you to be members of companies. You are a member of a committee. You're our advisers. We hope that you'll view yourselves as developing these principles for these guidances for industry. But you're not really industry and you're not really the FDA. We're hoping that you will bridge those two areas.

And we are hoping that if our ideas are misguided, you'll put us back on track. Of course, we hope that you too will enlarge your own thinking on the treatment of osteoarthritis and share that larger view with us.

CHAIRPERSON PETRI: Thank you. And now I'd like to ask Kent Johnson, a Medical Officer in the Division of Anti-Inflammatory, Analgesic and Ophthalmic Drug Products, to present the document to us.

DR. JOHNSON: I believe--does everyone have a copy of it, by the way? Hopefully most of you had it ahead of time so you have some flavor as to what this is about. I think Mike has set the stage nicely so I won't belabor the sort of overall philosophy of this meeting, and I am pleased with everybody that has attended. I think amongst the audience we have a huge wealth of experience with designing and conducting and analyzing trials in osteoarthritis. So that is really what we want here. We want vigorous and frequent feedback. So we are going to structure the meeting that way and have lots of periods for give and take. We don't have a lot of microphones in the back but make use--we can probably even hear you without microphones, as a matter of fact--okay.

I think most of you are probably familiar with this guidance process that we've gone through with

rheumatoid arthritis, and the intent is to try to do something analogous with osteoarthritis. Obviously some of the issues are different. Some of the issues are the same. The pure analytic--there are certain analysis issues and trial design issues that are really identical with osteoarthritis as in rheumatoid arthritis or in assessing any other chronic symptomatic disease, and in some ways we really don't have to revisit those issues.

But the goal today really was quite fundamental, and that is to try to get a discussion regarding what should be valid, what should be sort of robust claims for a drug that one asserts is effective in the signs and symptoms of osteoarthritis. We don't want to have--I don't think it would be useful at this point to have very detailed arguments about, you know, one assessment method versus another. There are these sort of metrology issues that are important, but we're looking at larger landscape, I think, today. As you can see from this three to four page handout that you have, we have preliminarily, and I emphasize preliminarily, come up with actually a separation of a number of claims.

One could actually argue at the outset that you shouldn't separate claims or at least short to moderate-term symptomatic claims. Maybe they all should be lumped

together. That's a discussion we should have. We have separated them: pain, function, structure, durability. The Omeract Group, which I think most people here are familiar with, had sort of a consensus driving meeting in Australia about two years ago now where there was virtual unanimous agreement that certain domains should be assessed in all OA trials. And those domains were pain, function, some sort of overall patient global, and for trials of a duration of a year or more X-ray.

So there is some parallelism here between what we've done and what the Omeract Group has asserted, but there is a difference, too. Those were the domains that should be measured in every trial, and what we're talking about are the bottom line for the companies and the bottom line for our job of trying to assess data, and that is what claim one is pursuing and how it's going to be written in the label. One big issue will be how does one assess other important domains if you're going after just one particular domain? If you are going after a pain claim, what has to happen with function or what has to not happen with function?

The structure debate is critical because of a number of new hypotheses in the development world that people hope will protect cartilage and hence protect the

structure and hence, at least in the longer term, alleviate signs and symptoms.

The durability claim was really an attempt to set a higher hurdle, really quite analogous to what we did in rheumatoid arthritis. You can have dramatic drugs on the short to medium term, but the issue from a patient who has a 30 year disease or is looking forward to a 30 year disease or a 20 year disease is what's going to happen five, ten years down the pike.

And finally, delay in new OA development and delay in surgical joint replacement are, again, some sort of fantasies, people might argue. I'm not sure new OA development would be that tricky to design, but we should discuss that. Delay in joint replacement obviously brings up all kinds of non-clinical and non-medical possible confounders, and, you know, these are going to be challenging notions. But obviously the attraction of these two endpoints is that they are nice and clear cut.

And then we just put in another category there to encourage you, as Mike has just said, to not limit discussion to what is written. There may be major omissions in your mind and we need to know that. We need to hear that. So I think what we're going to do is sort of march through and have a discussion about each of the claims, an

open discussion with the audience, and then there are dimensions of trial analysis and how one integrates the evidence from various trials that really are often part and parcel with a discussion of the claims themselves so that we could try to get a sense of the overall landscape as to where you have to be to submit an NDA going after claims X, Y and Z.

And there are in your initial handout seven sort of fundamental questions. I made up a list of some more questions that pertain to each claim itself, and I am sure there are many that are not listed, too, that deserve to be addressed. So I am going to turn the mike back over to Michelle and we'll proceed.

CHAIRPERSON PETRI: Thank you. We're going to now begin the open public hearing. There are three registered participants in the open public hearing. We'll begin with them. The first is Dr. Steven Geis, the Executive Director for Clinical Research at Searle, who wanted to comment on standards for trials.

DR. GEIS: Thank you, Dr. Petri. I'm Steve Geis, Executive Director for Clinical Research at G.D. Searle, and my group has conducted several osteoarthritis trials over the past ten years, and the draft guidelines provide some

very important suggestions about the types of OA claims and the types of data that are needed to support these claims, and I appreciate the opportunity to just make a few brief comments about these guidances.

We believe that the claims for treating the signs and symptoms of osteoarthritis should require replicate studies of at least 12 weeks duration. At least one of these studies should include the knee as the signal joint and at least another study should include the knee as the signal joint. The studies should be placebo controlled and should include an active comparator to validate each study. The studies should be performed on patients who do have active disease or, if you will, should be in a flare state. And we do believe that adequate and thoroughly adequate dose response data should be obtained from these trials.

The primary measures of efficacy should include, as suggested in the guidances, the patient's global assessment of the arthritis, the patient's assessment of pain, the physician's global assessment of the arthritis as well as the WOMAC, and these should focus on the signal joint. However, as described in the guidance, secondary measures such as the nonsignal joint patient global assessment we believe should also be measured.

Now based on our data, we believe that the claim

for function should not be distinguished from pain and require six months duration of observation. In our experience, there has always been a correlation between pain and measures of function and the maximum effect on pain and function is always established in a 12-week study. We agree with the guidance that the claims for structural and durability changes should require at least 12 months observations.

Finally, the risks associated with any new agents in a new class of compound should be clearly described with specific clinical trials and clinical data. The new compounds that are on the horizon are the specific Cox 2 inhibitors. And we believe that the specific risks of these compounds requires certain types of studies. We think that replicate endoscopy trials using ulcers as an endpoint should be conducted. In these trials, at least two times the full therapeutic dose should be assessed. In addition, we think clinically significant GI endpoints, such as bleeding and perforation, should also be collected. And then also there should be specific studies on the platelet effects and the renal effects of these compounds with using doses of two times the full therapeutic dose should be used.

In summary, the efficacy and risks of any new agent from a new class of drugs should be taken into

consideration when considering the claims for that new compound. We suggest that guidelines for collecting data on the efficacy as well as the safety should be clearly put forth. Thank you for your attention.

CHAIRPERSON PETRI: Thank you. The next participant in the open public hearing is Dr. John Beary representing Procter & Gamble who wished to discuss linking claims to pain relief.

DR. BEARY: I'm Dr. John Beary from the Clinical Research Department at Procter & Gamble, and as mentioned I would like to raise the issue of linking the structure claim and pain and function claim. The context I'll put this in is a study of early to moderate primary osteoarthritis in a context where structurally you might be looking at ten to 15 year history from the identification of early OA to joint death, if you will, that bone on bone, total knee replacement situation.

And a question that might focus this as the group thinks about this today: would not a drug that significantly preserved joint space width and, of course, had a good clinical safety profile be valuable for early osteoarthritis treatment? And think about this again in the context of the early OA patients typically aren't taking their analgesic medications daily and regularly. The pain phenomenon tends

to be a wavering situation where there are good days and bad days, and you can find actually long periods where people are not taking analgesics in early OA.

Some factors I'd ask you to consider are these four. First reflects on our experience with RA where the treatment pyramid basically has been reversed in recent years so that there are earlier interventions realizing that if you don't save structure early, the game is over. Rheumatoid is an aggressive joint destroying disease. All the answers aren't there yet, but I think the new treatment paradigms show an interesting analog that it is wise to save structure early if you wish to preserve the function and get clinical benefit.

Secondly, if you are going to depend on a slow-acting drug which may or may not have analgesic benefit, will you get your patients to take a placebo over a two-year study? I think that would be very difficult in the conduct of clinical studies to do that if your drug might primarily have a structure effect. You would hope it would have others, but the time frame of two years may not be when you would see it. You might need five or six years as I've seen other commentators in some of the handout materials today.

Third, early OA research is going to be very

challenging to conduct and I would caution you to think about not making it so complex it can't be done. So some interesting things to ponder in the early days of these guidelines.

Finally, looking at the handouts today, I noted Nick Bellamy on page two, the second paragraph, brought up the same issue I did that it, in our view, would not be a good idea to link the pain effect to a structure effect in early primary OA. And I also note looking at the EMEA guidelines that were handed out today as well, at the bottom of page one, as they classify these drugs, start to think about them in studies, that structure modifying drugs may or may not have an independent effect on symptoms. And I think that captures my thought.

My final comment is a very brief one. Our experience in using the Likert scale has been much preferable to using the VAS. The VAS scale has not performed as well in the elderly and we'll send you written comments with references on that point. Thank you very much.

CHAIRPERSON PETRI: Thank you. The next participant in the open public hearing could not attend, and our Executive Secretary, Kathleen Reedy, will read the

response from Dimethaid Research & Company.

MS. REEDY: Dimethaid Research is a pharmaceutical company engaged in development and commercialization of a topical analgesic containing the NSAID diclofenac for the treatment of osteoarthritic pain.

We have most appreciated our ongoing dialogue with the Food and Drug Administration, particularly through Dr. Michael Weintraub, Acting Director of Division of Anti-Inflammatory, Analgesic, and Ophthalmic Drug Products. In a recent discussion, members of the staff encouraged us to present for consideration of the Arthritis Advisory Committee the following issues which have arisen out of recently completed and upcoming trials.

(a) What is the appropriate duration for an osteoarthritis clinical trial? We have been advised that FDA requires a minimum six-week trial duration. A review of international studies on osteoarthritis carried out between 1983 and 1994 (see Therapeutic Trials in Digital Osteoarthritis) suggests that four weeks is most adequate in general and in particular most studies on hand osteoarthritis have been done for a length of only two weeks.

(b) Would covariate analysis, or ANCOVA, which uses baseline as a covariate be better for data analysis

rather than the normal ANOVA, analysis of variance, method?

(c) Should analgesic use be allowed before final assessment? If not, it results in a larger discrepancy between per-protocol and intent-to-treat groups. If allowed, the same question of covariate analysis arises again with acetaminophen use as the covariate. Although intuition suggests that total analgesic use would go down when using an effective drug, many other factors including the use of acetaminophen for headaches, other joint pain, and other pains in general may confuse results.

Please consider the relevance and acceptability of these questions for consideration by the committee and inform us of any developments and procedures that we should be aware of to prepare for the upcoming issue.

Paul Varady, Regulatory Affairs Assistant.

CHAIRPERSON PETRI: Thank you. At this time, are there other industry representatives or individuals who would like to participate in the open public hearing? Seeing none, we are now going to move to the actual discussion of the document.

On our agenda, we have divided this up into the different claims. The first claim is pain, which is on page four of the xerox that you received this morning, at the top of the page, that the primary efficacy variable is any

validated pain scale; secondary endpoints are function and nonsignal joint patient global measurement. Trial duration should normally be at least three months or six to 12 weeks if there is a large body of similar drugs in the same class. Nonsteroidal anti-inflammatory drugs being the only current example.

For the discussions, I would like to welcome everyone in the audience to participate by coming up to one of the microphones and identifying yourself. Let me ask Dr. Johnson if there are some specific things he would like us to discuss as part of this?

DR. JOHNSON: Well, they have already been touched on actually be a number of the speakers, but there is always the issue of the duration of a trial and in some ways there is an argument to having a trial longer than it might necessarily take. I mean it may be that there are agents that we could show would work for a two-week period of time, over a two week period of time. I think we have not been uncomfortable in shrinking down the trial duration to six weeks or so or six weeks to 12 weeks with regards to drugs that we are very familiar with. But if you make a call about efficacy for a brand new drug over just a very brief period of time, there is much more uncertainty about the longer-term use of this drug both from an efficacy and a

safety point of view, which is why we have tended to try to add the qualification that if it is a new class of drugs that we need a little more a duration dimension to the trial.

And also in this discussion, we might as well bring it up right now. We have to address this fundamental issue of in what way should pain stand alone or should it stand alone? And what should be asked for function, let's say? Obviously you probably would not want to approve a drug that was very effective with regards to pain if function went down the tubes. If it were very effective yet function diminished a little bit, there may be an argument to doing it, especially if it were a safe drug. So this has to be brought up. In bringing that up, the implication is you get into the issue of what happens with your analysis, how do you assess this, are you going to have a primary measure and then some criteria that restrict how effective the secondary measure, in this case function, has to be, or do you put together some kind of algorithm that encaptures both on a by-patient basis, even though we don't have any data to drive the decision about these analyses right now?

So, you know, there are a lot of questions that come up. I had listed on this sheet should this stand alone? What co-success should be needed with regard to

function? How would one define no deterioration, if that is going to be your test regarding function? And should they by combined into a by-patient index or by-patient test? So I think we would be interested in discussion from the audience and the advisory committee with regard to these issues.

CHAIRPERSON PETRI: Why don't we start with whether pain should stand alone because I think that is something that will definitely lead to a lot of discussion. Dr. Abramson.

DR. ABRAMSON: I guess just to start the discussion I would prefer to see pain stand alone. I think you have to measure function, but I think as a separate indication. At least as I understand the question that is being posed to us, you really want to know if pain is improved and look at function as a separate indicator. One of the reasons is that because you are going to be treating people at different stages of their disease, if they have advanced disease you may not get a functional response because of the nature of just the destruction that is at that joint but you may get a good analgesic response. Therefore, the absence of the functional response shouldn't dimension the potential analgesic effect. So I think it would confound by trying, as I understand it, by trying to

link the two, but I think you have to look at function because I guess the worse case scenario, as in the old Indican hip, you might put somebody on the drug that will give them pain relief and then have an acceleration of the cartilage damage. So I think you need that as an outcome measure just to look at what is happening to the patient, but I think they should be disengaged in terms of the indications.

CHAIRPERSON PETRI: I think there are several situations that we can all envision where pain and function are going to be completely separate, and I think a good example would be analgesics that have nothing specifically to do with the function of the joint. So, for example, a narcotic or an anti-depressant might very well reduce the pain of osteoarthritis but one would not expect it to have very much effect on function.

The other is this issue with the signal joint. David Felson who couldn't be here wrote a nice letter summarizing some of his comments, and one of his examples is a patient in whom you fix one knee, for example, the signal joint, that the pain in that signal joint might go right down to zero, but if you do nothing about the other knee, so, for example, if this was a specific joint injection therapy, the function overall might not improve. So I will

sort of second what Dr. Abramson said. I don't see how we can link these two. I think they should be stand-alone claims. Dr. Madrid.

DR. FERNANDEZ-MADRID: I would agree with what you both said. However, it seems to me that linking pain and function may be different in early osteoarthritis and advanced osteoarthritis. I think most of your comments are applicable to more advanced osteoarthritis when there is deformity of a joint, osteophytes and impediment in motion that really will not change too much and we would do a disfavor if we link them both together.

However, in early disease, and I think this is one of the problems that this document has and the field has, the definition of the disease itself and the definition of early disease, early disease pain and function are intimately linked. That is in early disease, you can improve pain and the function of the joint will improve. So I think there should be a distinction depending on the selection of the patients, what type of patients are utilized in this study.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: I just want to say I agree with all that has been said. I think that we are sort of taking the academic view that these are separate domains or axes. But

they are hard to disentangle for the individual patient. I mean I think we can do it theoretically, but I think, in fact, it is often confounded and interrelated.

I think the other thing we should keep in mind is that this is, you know, what we're doing is a caricature of how we treat OA now, and it's conceivable that we have a new agent that is not going to address pain that could modify the agent and improve function because it prevents the onset of pain. I just want to expand the discussion, I guess, a little bit in that I don't think there is any question if you took the patient's point of view and asked them or asked what brings patients into the doctors' offices, pain is the overwhelming thing, but there are people who come in from certain cultures who come in complaining of a limp; they are not necessarily in pain. People with Heberden's and Bouchard's are mostly interested in the cosmesis issue, and I think these are all sort of patient oriented things, and if you're going to put pain up and give it its own little box, I'm not trying to put these other things in as equal weight, but I think they are part of the consideration if you're really trying to take the patient's point of view.

CHAIRPERSON PETRI: Are there other thoughts about this pain versus function?

DR. LIANG: Could I ask one question? What about

assistive devices? I mean you can reduce a lot of hip and knee pain with a stick. And it's probably better than an NSAID in many instances or to sit down. What do you do with that?

CHAIRPERSON PETRI: Well, in fact, let's enlarge upon that because the ACR now has treatment guidelines for hip and knee OA with basically serial additive therapy with acetaminophen and capsaicin, for example. And it's not clear to me in this document how that is going to be built in. Let me ask Dr. Johnson if he wanted to comment on that?

DR. JOHNSON: Well, you know, you have to have a control and you can't withhold known therapy, known effective therapy, and I guess most people think Tylenol is effective. I think the use of Tylenol in clinical trials is still a challenge. I looked up--there are two trials, two major ones that I have found. Maybe there are others. We may have to eventually address the issue as to whether Tylenol as an active control can stand alone without placebo to validate its assay, you know, the usual argument?

But the one was low versus high dose, ibuprofen versus Tylenol, which was Ken Brandt's study, which was not designed to show equivalence. And the other was that long-term cooperative clinic trial, two years, I think, of Naprosyn versus Tylenol, where the dropouts were monstrous.

So it's hard to know what conclusion to draw, but the issue is always going to come up in these trial designs as to how to account for co-therapy or background therapy? It's sort of like the methotrexate situation in rheumatoid arthritis.

DR. WEINTRAUB: I'm sure Dr. Johnson meant to say acetaminophen but--

[Laughter.]

DR. WEINTRAUB: Didn't you, Ken? Yeah, it just slipped out. This is an issue which runs across all of therapy, what to do with hygienic therapy, and I would like to pose to it to the statisticians: do we need to create strata, create different ways of making certain that the same number of use canes and capsaicin in both or all three or 17 arms of a study?

DR. TILLEY: I think you're asking a question that comes up in almost any clinical trial, which is what do you do with standard medical therapy, and I think that what I have seen in the trials that I've worked in is that the important thing is to define what standard medical therapy is so that you understand that you're collecting the information. I don't think you necessarily have to stratify, but I do think if there is something like in this situation like acetaminophen that is standard therapy, you might want to try to see that your patients are getting the

same dose. So I think it's really agent specific and also somewhat specific to the time frame in which you're doing it and what the standard is at that time.

But I don't think that you can mandate the background very well. I mean I think patients themselves, things that they can get over the counter, they're going to do. So you really just need to know what's going on.

CHAIRPERSON PETRI: Dr. White.

DR. WHITE: I agree with the comments that have gone on, too. I was particularly struck by the problems where you would like to link pain and structure in early OA but maybe wouldn't like to--or pain and function with early OA--but you perhaps wouldn't like to lump them in studies of late OA. But it would seem to me that unless we are comfortable being able to distinguish those two groups of patients, that even though that would be desirable, it is not feasible. So that probably the pain should stand alone.

CHAIRPERSON PETRI: Other--oh, Dr. Johnson had a comment.

DR. JOHNSON: Is there anybody in the audience who was part of the European document? Maxime, were you involved with--do you have any--would you like to make a comment regarding your thoughts on this linking or non-linking.

DR. DOUGADOS: I have to introduce myself first because I am not involved in this document. I am Maxime Dougados. I am here as an individual, but I am a member of the Osteoarthritis Research Society and I also chaired the GREES section of Osteoarthritis and we published recommendations for conducting clinical trials in osteoarthritis last year.

Considering the discussion you are having during this meeting, I have two comments at this stage. The first one, just to remind you, that what we have done in the long discussion we had that to clearly differentiate between symptom modifying drugs and structure modifying drugs. Within the sub-group of symptom modifying drugs, we didn't discuss whether or not we have to differentiate pain and functional impairment. But what we have done, we have discussed the possibility by giving a symptom-modifying drug at the same time we may observe a deleterious effect on structure. In other words, if you kill the pain, the patient is painless, and therefore he will walk a number of hours during the day and at the end they will be a deterioration of the structure.

That is a reason why within the GREES, but also from the recommendation of the European Agency the claim at this stage is symptom modifying drug without deleterious

effect of structure. That is you want to claim on symptoms, you need also in the dossier to give some information concerning the lack of deleterious effect on structure. And this is not discussed in the manuscript that you propose, and that is the reason why, finally, I agree to differentiate between pain and function because function impairment at the end and structure will be the same. In other words, we can kill the pain within one hour and if you take the painkiller during two years, perhaps you would have a deterioration of the structure. So you have to be aware of that.

I am aware of a lot of difficulty in developing a new drug such as a painkiller if you write down this sentence: that is without deleterious effect of structure. that means that you need to bring in the dossier not only the demonstration of the improvement in pain but also the demonstration of the lack of deterioration of structure. That was the main point we had discussed during several hours, and the conclusion was that this one you have seen that in the manuscript coming from the European Agency, symptom modifying drug, whatever satisfaction but without deleterious effect of structure. But we didn't discuss the differentiation between pain and functional impairment.

As an individual point of view, I agree with you

to differentiate pain and functional impairment. Even I agree that at this time in clinical trials in patients with painful disease, with the tools we are using, such as Lequesne index or the WOMAC, there is a very high correlation between pain and functional impairment, but what will appear if you compare the changes in the functional part of the WOMAC between month three and month 24 in patients taking a painkiller? So to answer your question, yes, I do agree to differentiate pain and functional impairment, but please discuss the possibility of if someone wants to have a claim of chronic intake of painkiller, what about the possibility of structural deterioration?

CHAIRPERSON PETRI: Dr. Siegel?

DR. SIEGEL: In your document, did you have a recommendation about how long the study should be to rule out less significant impairment in function and structure?

DR. DOUGADOS: In function or structure?

DR. SIEGEL: Sorry. In structure.

DR. DOUGADOS: In structure, at least one year. And the more and more we conduct epidemiological studies, the more we reach the conclusion that if we focus in some population of patients, probably it's possible to see something changing within one year with the problem of the general--that is it is possible to see a change within one

year in the very specific subgroup of patients. So probably you would get some answer within this specific subgroup of patients. The problem will be is it possible to generalize to the world population of osteoarthritis, but within one year it is possible, at least one year.

CHAIRPERSON PETRI: May I ask at one year, do you mean by X-ray?

DR. DOUGADOS: Yeah, X-ray. Any well validated method permitting to evaluate the structure. At this time, we have three possibilities. The first one, which is well validated and simple, is X-ray. The second one which is more invasive is arthroscopy. And the third one, which is emerging but we are still waiting from research from longitudinal studies is MRI.

DR. JOHNSON: No drug would be approved for osteoarthritis if it didn't have one year of data that addressed the lack of structural deterioration?

DR. DOUGADOS: If the claim is chronic intake, no.

DR. JOHNSON: Okay. Yeah.

DR. DOUGADOS: That's a problem of the claim also. Do you have a claim for acute pain or chronic pain?

DR. JOHNSON: But for one year of fast-acting drugs, you would accept a three-month study and you would ignore the issue of long-term structure or at least in terms

of control data?

DR. DOUGADOS: That is still under discussion. Next week we will have a discussion with the drug company because for the demonstration of efficacy, it's quite easy within three months to demonstration such efficacy. The problem is the second part of the sentence: without deleterious effect of structure. And for this you need disease trial of at least one year duration.

CHAIRPERSON PETRI: Thank you. This has obviously complicated our discussion and it's opened up two issues. I think one that we haven't discussed is duration of trials for efficacy, whether it should be three months or less, and now we've raised this possibility of what should be required for this claim? Should it also require co-success meaning no detriment in structure at one year? Dr. White.

DR. WHITE: Regarding the latter point, Michelle, I was just sitting over here thinking that sort of a requirement imposes our judgment that that's what the patients would like. I think that there may be patients who would be delighted to get of rid of pain if it's significant even if it meant some decline in structure.

DR. JOHNSON: Yes, I think that perception has to be entertained, and we've thought about that a little bit. I think Jeff and I have gone back and forth on this, too.

If you had a drug that's dramatic with regards to pain, you might even accept some deterioration in function a little bit. I don't know. It may depend on the individual basis, but if--I think the issue, and it would be interesting to hear some of the statisticians either at the table or in the audience address this--how do you statistically or quantitatively define no deterioration or just a little bit of deterioration. Marlene, did you ever deal with that issue?

DR. EGGER: For osteoarthritis I can't give you a good validated measure. It certainly is an issue.

DR. JOHNSON: Let's assume that there is a good validated measure. How statistically, I mean if you had, let's say, just one clinical trial, and you're going to win by your primary, which is pain, but you don't want to deteriorate by function, is there sort of a straightforward statistically that you could quantitate that?

DR. EGGER: Well, you certainly can analyze the variables separately. If you're talking about a composite variable, that could be complicated.

DR. JOHNSON: No, just two separate variables. Let's say you've got a pain VAS and a function VAS, you know, both of them ten centimeter lines, so you've got nice clean data and you don't have any dropouts.

[Laughter.]

DR. EGGER: You know I don't think that's problematic. I have a little hesitation about whether we ought to penalize the drug for something that the patient is doing, but that's an issue, that's a recurring issue in studies of rheumatic disease. When the patient starts to feel a little bit better, what do they do? Do they go off another drug? Do they walk themselves till their disease is worse? I think that's a very difficult issue and it's philosophically more troubling when the statistics are involved.

DR. TILLEY: Also, if you're truly asking them to do something that you would call non-inferiority in terms of deterioration, you're going to impose some sample size problems for the sponsor, I think, because the sample size for the pain question may be much smaller than the sample size for non-inferiority in terms of deterioration. So I think you need to think carefully about what you want to do.

CHAIRPERSON PETRI: Let me ask you, Dr. Tilley, wouldn't that then be an equivalence?

DR. TILLEY: Yes.

CHAIRPERSON PETRI: So the sample sizes might, in fact, be huge?

DR. TILLEY: Well, potentially. I mean it depends

on the rates, although some of the newer approaches, the approaches to sample size, make it usually definitely larger than an efficacy trial.

DR. JOHNSON: But if your equivalency test allows for a pretty large window to be ignored, you know, that will help; right?

DR. TILLEY: That would help.

DR. JOHNSON: That will make your sample size smaller.

DR. TILLEY: But then you're back in the clinical domain and not the statistical.

DR. JOHNSON: Right. But I think there is a clinical question here as was brought up by Barbara. You know if a patient does dramatic vis-a-vis pain, they may tolerate a little deterioration of function.

DR. TILLEY: Yes, so I think that, again, I think you have to look very carefully at what you're asking when you make these kinds of linkages in terms of both the statistical and the clinical question.

CHAIRPERSON PETRI: Dr. Moreland, did you have a comment?

DR. MORELAND: Yeah. I'd like to comment. I think when combining the claim of pain with the change in deterioration of structure has several major problems from

the logistic of standardizing those techniques at a clinical investigative site. I know what's happening now with MMP inhibitors and being able to standardize radiology technicians and so forth is critical for MMP inhibitors but for another type of pain medicine, be it a nonsteroidal or Cox inhibitor, I think the numbers of patients in the standardization are going to really limit then the development of a drug. And I would offer a little different opinion than needing to add a claim about not showing deterioration as part of that. I would keep it alone just from the logistics of some of those issues.

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: Yeah. I would second what Dr. Moreland said. I think the deterioration of function is a critical issue that needs to get looked at, but it really should be separated from whether a drug is indicated or is effective in relieving pain in these patients. Because I think the issue of validating those outcomes at one year is still too early to create a linkage there.

DR. JOHNSON: So are the two of you saying you should let function just sort of fall out like you would let, you know, GI tolerance fall out.

CHAIRPERSON PETRI: Structure.

DR. MORELAND: I would like structure fall.

DR. JOHNSON: Not function though?

DR. ABRAMSON: No, we moved on to structure.

DR. JOHNSON: Okay. And then you would assess--

CHAIRPERSON PETRI: Wait, wait, wait. We've already separated pain and function. And now I think what we're saying is we want to separate pain and structure as well.

DR. JOHNSON: Yeah.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: I want to third the comments, but I also want to introduce some caution about this one year magic duration because I think that we're talking about the radiographic assessment in a model which has only been studied once as far as I know to show that you can enrich a sample for people who will have enough change over a year. And that's if someone, a woman has one knee, you know within the next year the other knee is going to have progressive cartilage change. That's one study in the UK. It's the premise on which the doxycycline study is being done, but I think that whether arthroscopy or other serological markers would be better or worse or whether that will even hold up in a trial, I don't think we know. We haven't really had much experience, real experience, and I think that doxycycline is something that we need to watch, but the one

year thing I think is something from a tablet in the sky, but I don't know if we can really put much faith in it, and I wouldn't put it in print certainly at this point.

CHAIRPERSON PETRI: Why don't we move on and discuss duration. Excuse me. Dr. Madrid had a comment.

DR. FERNANDEZ-MADRID: I want to raise another issue that it is a little bit separate but if we are talking about separation of pain and structure and we are considering drugs that potentially may lead to deterioration of structure, I think we may raise some ethical issues and whether joint protection could not be included in the study in some way?

CHAIRPERSON PETRI: Yes, Dr. Luthra.

DR. LUTHRA: I just wanted to make a comment about the issue of pain and function and deterioration of the joint. Right now, the best parameter that we have to measure deterioration of the joint is the X-ray. I think this is part of the reason we are having difficulty of combining the two or separating the two, but if we had serological markers which predictably predicted that the outcome is going to be poor and those could be influenced by the therapeutic agent that we are using, then I think we could combine the two very easily. But right now we're kind of looking at pain which is patient's response to an

analgesic. We talk about if the pain response is appropriate and the patient feels better, the activity level is going to increase, and at least the current knowledge is that increased activity under those circumstances will deteriorate the joint, but we do not have the biochemical measures as yet which would help us follow these joints and see can we measure deterioration or can we show that indeed this particular agent gives them pain relief and also protects from deterioration. So the whole question really comes up as how do we measure functional deterioration and is this just going to be symptomatic? Is it going to be X-rays? Those issues will really influence how you're going to put these two things together or separate them.

CHAIRPERSON PETRI: Okay. Now I did want us to discuss duration since this has come up in the open hearing and also in some of the letters we received. It is suggested here three months for a new drug. Are there comments? Dr. Harris?

DR. HARRIS: I believe that some of this, you know, one has to think as a practicing clinician in a sense, and in terms of somebody with OA and their treatment, certainly the practicing clinician, a patient with OA to whom one gives an nonsteroidal or any other analgesic, really one expects to get a pain response certainly within a

few weeks. One may then ask the question then why not two weeks or three weeks? But my belief is three months is really just about adequate, 12 weeks, because on the one hand one may have early response, but you certainly do want to keep the agent long enough so that if there are any untoward effects that might occur. Perhaps three months might be a reasonable period in which to think of looking for those.

Certainly clinically if a drug does not work within three to four weeks, it probably isn't going to do very much. So certainly a three months period of time, in my view, enables one to establish pain relief and hopefully it gives you enough time to assess whether or not there are any early untoward effects that might occur.

CHAIRPERSON PETRI: Other comments, thoughts about duration? Let me ask the audience for those who wish to come up to discuss both the separation of pain, function and structure and duration.

DR. BEARY: Thank you. John Beary, Procter & Gamble Clinical Research. The duration issue, are you addressing to all three aspects at this point?

CHAIRPERSON PETRI: No, just for pain.

DR. BEARY: Just for pain. Okay. My comment--

CHAIRPERSON PETRI: And we have lumped, you know,

we haven't separated out hip, knee, digital.

DR. BEARY: Right. My comment was more to the duration of structure claims so I will hold that for later.

CHAIRPERSON PETRI: Other comments from the audience? Please come to the microphone.

DR. SHAINHOUSE: I'm Dr. Shainhouse, the Medical Director of Dimethaid. In fact, that was our letter which was read into the ledger earlier on. And we're particularly interested in the assessment of pain. What I'd like to understand is if we do separate pain from function, and I just want to make one point, we certainly agree that pain and function are clearly separable or to be separated in assessment, bearing in mind that every validated method for assessing pain really assesses pain through function. We're simply calling these functions so basic as to be part of daily life. Sleeping is a function. Sitting in a chair for someone who has a hip discomfort is a function. Walking, which is part of most simple pain assessment, is a function. So it's all a degree of definition of function.

Nonetheless, the detailed distinctions or questions in function can remain separate, and that makes sense to us. If we also appreciate that pain and structure are to be separated, and given that we understand that in this disease there is a very poor correlation between

radiological structure and pain, particularly in the earlier stages, it's very conceivable that the patient will be quite happy to have pain relief and indicate pain relief on a study or in clinical practice while there is a natural history occurring, namely minor deterioration of the structure before the very eyes of the physician if he or she chooses to repeat the films at intervals. Does that mean that will deny pain relief to our patients because there could be some deterioration of function over time?

The other obvious answer is in different diseases, and I always look to different disease models, it's like telling the patient with angina just stay in bed. Well, that is not what coronary bypass surgery the most commonly performed operation in America. So one has to put that into a perspective. Those are the essential features that I wanted to stress again with pain structure and function. When we get down to the length of a study on pain, if we separated out these issues, particularly pain and structure, pain and function, what reason then do we have for wanting to study pain beyond the time which is necessary to determine that the pain relief has occurred so that if individuals or if a study can give a statistically valid response within two weeks, three weeks, four weeks, that pain has been relieved, what reason do we have in our mind

for saying I don't accept this, I need to look longer? Are we talking about narcotic where we feel that the concept of tolerance exists? I don't recognize that for NSAIDs. I don't recognize that for acetaminophen. Are we talking about going for a longer and longer study where the natural variation of the disease will come into play? That will just make it more difficult to demonstrate a valid response, particularly in early--what are the reasons?

CHAIRPERSON PETRI: Let's start to address some of your comments. I think Dr. Harris addressed one. The reason that the committee felt comfortable with a three-month duration of the trials was for safety issues. But I think also we're thinking that there are going to be new drugs for new mechanisms of action developed and for those there might be an issue of tolerance and durability. So this document is not just for the known. It's hopefully going to prepare us for the future as well. But let me open up for the committee responses to the remarks just made? Dr. Moreland?

DR. MORELAND: I would like to agree. I think if one looked at rheumatoid arthritis as an example, if we looked at deterioration of radiographs and withheld a nonsteroidal based on that, we wouldn't be giving any nonsteroidals. I think the point ties in very nicely with

some of the other comments. I really think that when you're looking at such a heterogeneous group of OA patients, whether they're obese and they have one knee involved and they don't have the other knee involved, to tie in structure and function with pain relief is going to be very difficult, and I would keep them separate, and I agree with the comments you made.

CHAIRPERSON PETRI: Another comment from the audience. Please come to the microphone.

DR. DOUGADOS: Yes. Maxime Dougados. If you focus discussion of the efficacy of the drug and the duration of the treatment related to the efficacy of the drug from both known and unknown drugs, I just would like to emphasize the fact that for the known drugs such as analgesics and nonsteroidal anti-inflammatory drugs, the duration of days, of one week is sufficient to demonstrate the efficacy, but I just would like to emphasize the need--I am a rheumatologist--to know not only statistical significance within two drugs--that is it works--yes or no--but we need also as a clinician to know the onset of action, when the onset of action is occurring and also when the plateau of efficacy is reached.

In other words, with a new drug such as we have in the European countries some drugs which are the onset of

action occurs within four or six weeks, but the plateau of efficacy is only reached after four months. In other words, it's possible to demonstrate statistically significant difference between the placebo and the study drug within three or four months, but at least in my own experience, I missed the information concerning when the plateau of efficacy is reached, and in our daily practice that is very important to inform our patient. You take this pill, but you will have to take this pill at least and you give the number of weeks or months that is a time when the plateau of efficacy is reached. And, therefore, I think it is the reason why in the European document, there is no figure, but the sentence was related to what you think about the mechanism of action of the compound.

CHAIRPERSON PETRI: Thank you. Next comment from the audience.

DR. HOLFORD: Yeah. My name is Nick Holford, and I'm a visiting professor at the Center for Drug Development Science at Georgetown University. I'm a clinical pharmacologist, and I have an interest in the modeling of diseases and drug action. And I'd like to make the comment that in discussions I think you should distinguish two perspectives. One perspective I will call the regulatory perspective which says we need to answer the question does

the drug work? The answer is yes, no, or sometimes maybe.

The alternative perspective, and I'll take this as the perspective of the patient and the physician, is what do I know about the drug so I can use it effectively? And so I'd like to echo the remarks of the previous speaker. We do need to understand the time course of the onset of drug action, how long does it take to reach its peak, does it fade away, do we develop tolerance or not, and can we disentangle that from the time course of progression of the disease itself. So in relation to remarks about study trial design, I would suggest it is important that we have, first of all, repeated measurements of whatever it is you're looking at frequently enough that you can describe the time course of the drug action and the disease state and it's also got to be long enough that you can observe the onset to a peak effect, and also find out whether or not the drug effect is disappearing or not. Thank you.

CHAIRPERSON PETRI: Is there another comment from the audience?

MR. LIPMAN: Yes. My name is Bruce Lipman. I work at Pfizer Central Research. Two comments or perhaps three. One is I would like some clarification with respect to duration. For a chronic use drug, I would assume that we would be obliged to treat under ICH guidelines for chronic

use drugs in duration at least for safety evaluation, and I'm wondering how the committee would view stopping looking at efficacy at some point and only looking at safety? I don't really see that happening. And so I'm wondering if it's kind of a moot point anyway.

Secondly, with respect to linking structure and function with pain, Maxime Dougados mentioned that if you have an analgesic drug, that perhaps if it were a good analgesic drug, patients with osteoarthritis would use their joints more and deteriorate more. And this suggests that, in fact, if you required no progression of damage to an osteoarthritis for a drug for pain that it would actually have to be a structure modifying drug as well, and so that would really complicate matters. Plus I agree totally with Dr. Moreland and Abramson that just the doability of clinical trials in multi-center studies when you're trying to evaluate effects on pain, to have to also coordinate everything to make sure that measurements of joint space width or something like that are interpretable would be a logistical problem, very expensive, and make every osteoarthritis study into a study of structure modifying agent which would really hinder research in the area.

I know that our company would think that if we had to invest like that in every drug that we were developing

for osteoarthritis for symptoms, it would mean we would probably develop analgesics for something else.

CHAIRPERSON PETRI: Let me reassure you right now that the consensus from both the committee and the audience has been that we need to keep these claims separate, but let me ask Dr. Johnson to address the specific issue about length of study for safety.

DR. JOHNSON: Yeah, well, the ICH guidelines pertain to safety, and I think we have to realize that any efficacy inference from one year is going to be highly uncertain unless you have a control, and if you do have a control, it's still going to probably be highly uncertain because of dropouts. I mean this seems to be the fact of life. There may be sophisticated ways we can deal with this, but the few nonsteroidal trials that I've seen even going six months, let alone one year, you just have so many dropouts.

We have not made a call about what happens to structure. There could be an argument obviously that if at the one year point in your ongoing follow-up, all the knees went down the tubes and they all needed joint replacement, then you could argue that this is a safety problem, and you have to address it. So it will come back into the equation, into the risk-benefit equation. But to get it there

formally and systematically requires these long-term control trials, which you know with certain active controls might be more feasible. But the point Maxime brought up about slow-acting pain relievers or pain/function relievers, we didn't explicitly address. We probably will have to because, you know, you're absolutely right. You need longer trial durations. If you did have a drug that took six months to start to work, you could actually show that in a trial, I think, if your trial was large enough even if the maximal effect or the plateauing didn't occur until say nine months or something like that. It's going to be hard to know how to make a call. I think there is going to need to be individualization of these things.

We put in these time figures as minimum figures, and obviously if you've got a slow-acting drug that doesn't kick in for at least three months, then a three month trial is not going to do anything.

CHAIRPERSON PETRI: Dr. White?

DR. WHITE: Again comments related to duration. it seems to me the numbers were put in based on analgesics because you thought that would be long enough to address the issues of onset, efficacy, tolerance, even though that apparently is not an issue. But that presumes that you're going to be just dealing in the future more or less with one

type of drug and they'll all follow this model. Why not just say what you would like and let the duration be based upon the drug, that you'd like to have the trial long enough to demonstrate onset of action, plateau of efficacy, tolerance, and let it be designed according to each drug?

DR. JOHNSON: Well, I think there is an argument to having a cutoff. It's bound to be an arbitrary cutoff. In the RA document, there was just too much nervousness about these agents getting tossed into people, and they work very dramatically. They'll work in a few weeks time. But we said, you know, unless there is a major heritage of experience with drugs of that class, which even in the RA world is essentially still just the nonsteroidals, you have to out six months. We backtracked, I think we're trying to be sensitive to, you know, these developmental dynamics that obviously exist here, and you know, I guess it would be incumbent upon the sponsor to have an adequate package so that we were comfortable that even with no control data beyond three months--let's say you do your three-month trials, and you win by pain and you don't lose by function, and you've got open extension out to a year, I think the bottom line is still going to have to be that, you know, in the risk-benefit analysis, it looks acceptable, but that's all judgmental, which makes people uncomfortable.

CHAIRPERSON PETRI: Yes, Dr. Tilley.

DR. TILLEY: Well, I've heard a little bit of lack of clarity about what we're talking about with no deterioration. I heard Kent talking about the problems with the comparison group and then I've heard other speakers saying, well, we can't expect these patients not to deteriorate because the patient not on the medication might be deteriorating. So I think we need to think carefully about whether we're talking about no deterioration in the context of the patients on the treatment or no deterioration in the context of some comparison, and then we have to think about what that comparison might be, given the difficulties in keeping some other group off whatever for a long period of time to make that comparison. So I think it becomes increasingly difficult and really does require a lot of precision if you're going to put that into the claim.

CHAIRPERSON PETRI: Are there other comments from the committee? Dr. Liang?

DR. LIANG: This is just a--I mean I know we're all trying to get efficient trials and keep the costs down, and I obviously want to see new agents developed for this, but I think we're sort of taking our RA experience into OA. I mean most of these measures are questionnaire based and could be done over the phone. I don't think there is any

need to examine these patients, you know, repeatedly. They don't really add anything to it, and I would rather see the data or the money spent in that structural question irrespective of what we decide about the duration recommendation.

CHAIRPERSON PETRI: Comments about doing all these studies on the phone?

DR. CALLAHAN: Or you could do them by mail.

DR. LIANG: Or whatever. You don't have to bring them in.

DR. CALLAHAN: Yeah. No, you don't have to bring them in for the pain or the function.

DR. LIANG: The phone call may be actually beneficial.

CHAIRPERSON PETRI: So we're just talking about the pain claim?

DR. LIANG: No, no. Function, too. I mean all of these are basically a person's report of whatever.

CHAIRPERSON PETRI: But, Dr. Liang, don't we also want to know, for example, if there is a joint effusion which is increasing?

DR. LIANG: I mean you might, but I'd give that up for the report because I think we all, you know, treat the person's complaints, not the effusion.

CHAIRPERSON PETRI: But in the development of a new drug, for example, if pain was decreasing and yet a joint effusion was increasing, wouldn't you want to know that?

DR. LIANG: You would.

CHAIRPERSON PETRI: I don't want you to put rheumatologists out of business here.

[Laughter.]

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: First, we give up the stethoscope, now the physical exam entirely. I would just make two points because OA is not RA. Therefore, when you continue trials beyond three months and commit to a year, some doctors really don't treat patients with OA that way with daily nonsteroidal drugs, at least, for that duration. So I think the three months is a benchmark largely for safety issues. When you get much beyond that, it becomes impractical almost to run some of these studies in terms of patient recruitment.

The other point I would make is just as a side point that some of these drugs may have more than one mode of action so, therefore, you may see, for example, in the Cox 2 inhibitors some analgesic effect early on, but there may be effects on structure and metalaprotenase activation

that won't be seen for two or three months. So when we think about evaluating these drugs, we have to keep the fact that they may do separate things in mind as we design studies.

CHAIRPERSON PETRI: Let me ask--oh, I'm sorry. There's another comment. Dr. Strand, if you would come to the microphone, please.

DR. STRAND: Hi. I'm just a little puzzled, and the puzzle that I have is even as a rheumatologist or even as a patient, I don't think I know how to separate pain and function. And I don't see how we really can because the pain will determine how much function you are going to do. And in that sense, you know, if people are continuing to have limitations, then they may not have as much pain, but, in fact, you haven't really given them something that is a meaningful relief. So I don't know. I can understand how we can say that you should have a product that improves or decreases pain, but I don't know how to separate that in the context of it means that they have less pain because they are now not doing anything, sitting, for instance. Have we really offered any kind of benefit? So to me whether we separate structure or not still makes a lot of sense, but I don't know how to separate pain and function.

CHAIRPERSON PETRI: Responses? Dr. Liang?

DR. LIANG: Well, I completely agree with you, Vivica. In fact, that's why I said I think we say we're going to separate it, but in fact I think what we measure is a mixture. I think the pain report is most valid when it's put in the context of how much symptom did you have when you walked, a certain kind of activity. But I think we're joking, you know we're kidding ourselves if we think we can keep it separate, and I think sometimes for patients it's very hard to distinguish the two as well.

CHAIRPERSON PETRI: Dr. Johnson.

DR. JOHNSON: Let me ask the metrologists or the statisticians in the crowd, if you have two variables that are sort of codependent, like Vivica was saying, if your pain goes up, your function goes down, if your function goes up, your pain goes down, could you make an argument that it's more sort of information efficient to not measure them separately and analyze them separately but to look at patient X and say what is the composite of pain/function in that patient, look at the next patient and make that same determination? In other words, a by-patient analysis. Charlie Goldsmith has written on this stuff trying to describe the scenarios where it's efficient and the scenarios where it's inefficient. I don't think he's here today, but I mean maybe that's what we're talking about in

essence here, and if it's true that it would be more information efficient to assess each patient individually and then add it all up and compare your two arms, then I think it sounds to me like the inevitable logical conclusion is combining the two claims into one and having it like the Europeans are proposing as symptom relief. Do any of the statisticians have thoughts on that?

DR. TILLEY: Well, I've been looking a lot at different approaches to multiple outcomes and the problem with these composite scores is getting something that's clinically acceptable and meaningful, and also there are issues of variability and distribution, and I know Charlie's work. He's been doing a lot of regression analyses, and they haven't been widely used, and I think it's because of the interpretability issues. So I think at this point, at this stage, that we still if we're talking about something for pain, that we may want to have function as a secondary outcome. But it isn't necessarily so that you add the two together and can do a smaller study. You know it depends on what the distributions are.

CHAIRPERSON PETRI: I'm still concerned about this issue of the signal joint versus the nonsignal joints. And injection therapy may relieve pain in the signal joint. Perhaps it will even improve function in the signal joint,

but the nonsignal joints may determine overall function such as ability to walk, get up from a chair. I still think that it makes sense to have these as separate claims, but let me again open it up for further discussion. Dr. Harris.

DR. HARRIS: I suspect it's been said before, but again function is to some degree, although it may be to a good degree, dependent on structural damage. And I think that's been said before. I mean surely patients may have functional limitations because their disease is really so advanced. The joint destruction may be such that they may have pain relief and not get functional improvement. My own bias is that because there are other factors that influence function that are separate from pain, that perhaps it's wise to separate pain and function--my view.

CHAIRPERSON PETRI: Dr. Pucino.

DR. PUCINO: Yeah. I agree with everything that has been said. You really have to separate the two, particularly if you're talking about early onset versus chronic.

CHAIRPERSON PETRI: This might be a good time to move on to the second claim of function. Again, that's on page four of your handout. The primary efficacy variable is any validated knee or hip OA function measurement. Secondary endpoints are pain improvement and nonsignal joint

patient global measurement. Trial duration should normally be at least six months. So I think that there are several things that we will want to discuss here. Perhaps we need to discuss the function measurements. We again have this issue of linking things to secondary endpoints, and the third issue would be trial duration. Why don't we start with function measurements. Let me ask Dr. Liang to maybe lead this discussion a little bit.

DR. LIANG: I want to badger the group again. What about the stick? You know what about lateral taping of the patella or patella femora. I mean we sort of dodged it with the pain, but the same thing, to sort of put it in your face.

CHAIRPERSON PETRI: So now you're going to hit us with the stick.

DR. LIANG: I'm not trying to say I know the answer, but I think you have to deal with it, at least treat it as a covariate. Collect the information but if you want to do the reductionist bit and make everyone use a stick, Tylenol, before you give them "x" drug, you know, those are two ways of dealing with it. But I think you can't put this template out there and let people game the comparison.

CHAIRPERSON PETRI: Well, I'm going to agree with you. These have to be part of the covariates or there has

some to be some stratification.

DR. LIANG: And do you use a stick? Do you hold the stick? Do you push down on the stick? 12 hours?

CHAIRPERSON PETRI: Other comments about functional measurements? Maybe I'll still put Dr. Liang on the spot. Do you want to discuss WOMAC?

DR. LIANG: Well, you know, there is the Tower of Babel. There are a zillion instruments for measuring function. It's a very personal thing, and it has to do with the person's aspirations, what they need to do and want to do, and you can have two people with identical X-rays, one who is really walking a mile and another one is sleeping all day, and I think this is sort of changing paradigm in clinical research that we want to know whether it fixes the X-ray but also whether it makes a difference to the patient. I think you have to collect it, but it's really sort of, you know, it's squishy. But it's better to have it than not. And I think all of the measures are fairly interchangeable. Anything that is published is pretty good. And the joint specific ones tend to be more sensitive, no surprise, because they have a lot more items than the generic ones, but the generic ones, you may need to do if you're going to try to collect information on the surrounding joints or comorbidity that might affect someone's ability to walk

like, you know, peripheral vascular disease. The generic instruments are also a hedge against missing symptom side effects that you don't measure actively because you haven't tried to look, you know, stringently for the side effects of drugs.

CHAIRPERSON PETRI: Are there site specific functional instruments for digital OA, spine OA?

DR. LIANG: There are hand function. Oh, spine OA, I think you could use several. There are a lot of options. I think there are 20 instruments for back pain, but they're all talking about structural pain, and I think they are all published and could be used. Nodal OA, I bet you could use a generalized hand function, and there are several of those, but I actually don't know there is a specific one for OA, you know, Heberden's and Bouchard's, but to me that's not so interesting clinically anyway. I mean I'm not looking for a new miracle agent for Heberden's nodes.

CHAIRPERSON PETRI: So all of you who are developing that miracle instrument, you're out of luck.

[Laughter.]

CHAIRPERSON PETRI: Other comments? Why don't we discuss this issue of tying function to secondary endpoints? We discussed this, of course, for pain, but now let's

consider it for function. Can you have a claim for function if there is no pain improvement? Dr. Madrid.

DR. FERNANDEZ-MADRID: I doubt it very much.

CHAIRPERSON PETRI: So we have one vote that you're going to have to have co-success. A function claim requires improvement in pain. Dr. Abramson.

DR. ABRAMSON: But I think it's the same argument we had before. It's unlikely that you will, but I think you still have to. If we felt we had to separate them in the analysis on the first piece of this, I still think we have to keep them separate on the second.

DR. CALLAHAN: Well, I would agree with that.

CHAIRPERSON PETRI: The keeping them separate? How can you convince a patient to take a drug that is going to improve function but worsen pain?

DR. LIANG: No pain, no gain.

DR. CALLAHAN: It might not worsen, it just might not dramatically improve or significantly improve.

CHAIRPERSON PETRI: So you're not going to require co-success, just no deterioration in pain? Dr. White is desperately looking for a microphone.

DR. WHITE: To be the devil's advocate, let's take your farmer who needs to get up and move around and move around a lot and, in fact, has a pretty high pain tolerance.

And he would really desperately like to get on and off his tractor and he could tolerate a little more pain if you could let him do that.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: Well, have we talked about stiffness? Because I think early disease and even people with established disease, there is a component of stiffness that is sort of the rate limiting step. And I think that those patients will say I don't have pain, doc. I just can't get out of bed and go as quickly, and so I can see--I'm making that up--but I could see someone improving in function and not saying something that helped him with his quote "pain" because that's not the way he expresses it, he or she.

CHAIRPERSON PETRI: Other comments?

DR. SCHWIETERMAN: Can I just comment on this?

CHAIRPERSON PETRI: Of course.

DR. SCHWIETERMAN: I don't think people in the agency dispute that if you're better off functionally and have the same amount of pain as before that you're better off overall. The question is in a clinical trial where you see only functional benefit, can you meaningful ascribe that functional benefit to the drug if there is no evidence of pain increase to a level of confidence that we're all comfortable with. Maybe you can. I don't know, but perhaps

that's a better way of looking at the question. It's not really a question whether being functional better is better--we all think that is--but can you do a trial where you see no difference in pain but see functional differences and be convinced that that's really because of the drug?

DR. ABRAMSON: In a hypothetical, we don't know always what the pain is due to in these patients. So hypothetically suppose you had a drug that restored cartilage function and increased the integrity of the cartilage, but since we don't always know where the pain element is coming from, the pain wasn't that much influenced. So at least in the hypothetical, you might see that kind of circumstance.

CHAIRPERSON PETRI: There's a comment from the audience.

DR. DOUGADOS: Yes, Maxime Dougados. I just want to come back to the problem of the requirement to dissociate clearly pain, function, what about stiffness, and you also discussed the problem of knee effusion, and after that we will discuss also the flare. There are a lot of domains to be evaluated in osteoarthritis, but, in fact, personally I have not the experience of a drug which is about to improve the functional tools we are using in daily practice. I am not speaking about the functional impairment of the patient.

I am speaking about the value we are obtaining when using the WOMAC or the Lequesne data. This is real life. And there is in my experience there is no drug which is able to improve the functional tool without improving the pain. And again I am afraid that if you are entering into the details, the patient and the physician will be completely confused. We have symptoms or we have structure. That is quite easy to understand.

And why are we afraid of the structure? Because we have the experience of long-term intake of Indomethacin in patient with OA. That is a patient taking Indomethacin may have an improvement in symptoms in the short-term. Short-term means six months. But at the same time there is structural deterioration, and, of course, the structural deterioration--of course, I am not sure--the structural deterioration within one year might be predictive of the long-term symptomatic deterioration defined by functional impairment. But I am afraid that if you enter into the details to clearly separate pain and function, someone will say what about stiffness? And then we say what about effusion? Do you want a specific claim for that also?

CHAIRPERSON PETRI: Let me ask for comments in response. We discussed this somewhat when we were discussing pain that to truly evaluate structure we thought

would probably require a year. Many of us thought that that might be daunting in terms of trial design.

DR. JOHNSON: Who knows the correlation in, let's say, the WOMAC, between the pain subscale scores and the function subscale score?

DR. LIANG: It's like unscrambling an egg. I mean the questions are functionally posed.

DR. JOHNSON: I'm sorry. What?

DR. LIANG: I mean most of the questions are posed in terms of a functional activity. And--

DR. JOHNSON: I know, but the way they are presented. I mean that or the Lequesne are the two ones that are being used, and he--you know, Nick has claimed that each subscale has been independently validated so somebody must have done an association study.

DR. LIANG: That's a statistical game in a way. That is sort of the Chronbach's alpha bit, which is that you try to see if items cluster, but that's a statistical thing. I mean I don't think that--

DR. JOHNSON: But does your score on the subscale, is there an association, a strong association between the score on the subscale, pain subscale, versus the function subscale?

CHAIRPERSON PETRI: Wouldn't one expect that? I

mean that's what we've been talking about that in many patients, especially the early patients, the two things should go together. We're more concerned about the late stage OA where these domains might not necessarily go together.

DR. JOHNSON: Well, I'm not sure they do go together that strongly in any group of patients, you know, which would be another argument for separating the two.

CHAIRPERSON PETRI: Dr. Strand had a comment.

DR. STRAND: I want to echo Maxime's comments because we've been talking about what to do with outcomes in OA studies for quite some time, probably even longer than even in RA. And I don't understand how, you know, we can really separate what a patient perceives in terms of pain or in function, and it makes in my mind more sense because everybody has different symptoms, different symptoms with different structural disease, different symptoms with expectations, different symptoms with different pain perceptions, that we consider looking at symptoms versus looking at structure. But how you can still in my mind separate pain and function is problematic even for how I see pain for my ownself. And I'm fearful here that we're all talking as if none of us have ever experienced OA and we're all getting to the age where we should have at least a few

tinges of it.

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: I would at least raise the question whether we're influenced by our experience of pain being related to function because of the drugs that we now have available for osteoarthritis which are predominantly analgesic drugs or nonsteroidals. If we're in an era of a new class of compounds coming out that may affect structure without affecting pain, we don't know if those drugs are going to unlink pain and function.

So my sense is that we should give the physician and the patient a clear tally this is what it does to pain, this is what it does to function, this is what it does to structure where we have the data, but not make a corporation or not make us have to interpret some linkage with a new class of drugs where we might not be able to predict if it may separate those outcomes.

CHAIRPERSON PETRI: Yes, Dr. Pucino.

DR. PUCINO: And that becomes more clinically relevant if we're talking combination therapy and we're using analgesics with something that modifies structure. So theoretically there is a reason to separate the two.

CHAIRPERSON PETRI: Dr. Schwieterman.

DR. SCHWIETERMAN: That's a point well taken, Dr.

Abramson. Would it be valuable in the set of guidelines to discuss priors that one might have before going into a trial with regard to the predictive mechanism of action--that's not the right word--the predicted benefits given the biology so that for an analgesic, for example, would you be equally comfortable with a functional claim separate from pain with a potential analgesic drug as you would with a cartilage rehabilitator or would you discriminate between the two?

DR. ABRAMSON: I think you have to--I'm not sure I fully understand the question. I think you have to discriminate the question that you're asking with any particular drug that you're trying to get an indication for. If you want to go in for function and structure, you'll have a certain kind of outcome, and it may take a year in the case of structure at a minimum. But I think you have to simply define your outcomes differently even for the same drug in terms of the indication you're trying to get but not limit the ability to get one indication because it doesn't do the other based on your a priori expectations.

DR. SCHWIETERMAN: I guess my concern is that I share your opinion that there may be drugs out there that improve function without pain, but to the extent that the clinical trial data even including the priors don't support separation of the two, it may be difficult for people to

really believe that something not previously billed as a particularly new clinical drug could, in fact, improve function. It may be an artifact of the trial, and that becomes very sticky with the subjective endpoint like function where you typically look at a whole range of other cooperative measures to substantiate that. And it gets tough when you say, well, we just need a function without anything else. But I see what you're saying though.

CHAIRPERSON PETRI: Dr. Madrid?

DR. FERNANDEZ-MADRID: I think although the correlation between pain and function is very high, I think an additional argument to keep them separate in gathering the data is the natural history of the disease that we are talking about. I think this is different in the knee, different in the hip, different in the hand, but there are flare-ups of the disease that may last a few days to a few weeks, and after the pain subsides we still have the function, and the function may be impaired by many, many factors which are operating differently in this multifactorial disease. So it seems to me that this is a reason to obtain the data separate for both.

CHAIRPERSON PETRI: Part of this claim is that duration of trials which is listed in the document is six months, and I wanted to bring that up for discussion. It

differs from the duration for pain which was three months. Comments about that, Dr. Callahan?

DR. CALLAHAN: Yeah, I would tend to agree with Dr. Felson's letter. I don't see a need to have the function measures at six months and the pain measures at three months. I think particularly the self-report function measures could be assessed for three months.

CHAIRPERSON PETRI: Dr. Harris.

DR. HARRIS: I was just going to ask why six months? I mean what was the thinking behind the six months?

CHAIRPERSON PETRI: Well, Dr. Johnson.

DR. JOHNSON: We wanted one of each time category.

[Laughter.]

DR. JOHNSON: No, I don't think there is a good answer to that. I think some of us had the sense that function, and I think this is a little bit what Bill was implying, that somehow function is a little more illusive or maybe a little slower to manifest itself or this or that compared to pain. I mean, you know, pain is very straightforward, and on a zero to ten scale, boom, you've got it for pain. So, you know, but behind all this is what durability dimension do we want for every claim, and we didn't necessarily perceive that they should be identical in essence. But this is wide open for discussion obviously.

DR. HARRIS: Can I make--you know, in my mind pain is relatively clean in terms of measurable response and, you know, it's pain improved, pain not improved. In terms of function, it's a little messier in my mind. I mean functional improvement could be indeed from pain relief, and indeed one might see an early response, but clearly what one is trying to get at here is that there may be other factors presumably structural or something that might, in fact, result in an improvement in function. And I guess the difficulty with respect to time is my whole difficulty with understanding, you know, function itself and just many, many more variables in terms of determining functional outcome. So six months, a year, three weeks, it probably in my mind depends on, you know, what are the various factors that contribute to functional impairment and, you know, any of those could be affected in any particular way. So it gets a little fuzzy and that's why for me time is a little fuzzy.

CHAIRPERSON PETRI: I'm sure there must be comments from the audience from our industry representatives? Dr. Geis from Searle.

DR. GEIS: Thank you. I'm still intrigued with the separation of pain and function. We have literally studied thousands of OA patients, and we collect the standard measures of pain on VAS and then the WOMAC and

other measures of function. In measures of central tendency as well a studying specific patient profiles, I have never seen it separated. Now it may be that we are studying short-term 12 week studies in patients who don't have, you know, very advanced disease, but the average patient has usually had a diagnosis for at least ten years, and we just keep seeing this repeated pattern that you don't see the two separate.

So I question how would you do an analysis of the data that says I really have a drug here that only works for pain and doesn't really affect function or vice versa? I just have never seen data look that way and I don't know if you could give us some guidance as to any studies or anything or types of patients that we really could look at it so we know what does it look like when you have a functional effect but not an effect on pain or vice versa?

CHAIRPERSON PETRI: Dr. Liang wanted to start the discussion.

DR. LIANG: Yeah. I mean there is data on this. I mean I think most of us would agree that the single most effective thing we can do for someone with an end-stage knee is put a new knee in it.

CHAIRPERSON PETRI: I thought you were going to say weight loss.

[Laughter.]

DR. LIANG: Weight loss. No. But if you look at a cohort of people who have had knee replacements and follow them up serially with any of the measures of function, and I should also make the distinction, and you should in your document, that you're talking about physical function, and you're going to have to deal with whether it's generalized function or function of the signal joint, but be that as it may, you can see that the trajectory of function is it improves after three, but it gets better after six, and there is varying sensitivity to that effect by which measure you use.

And I think that that actually makes some intuitive sense is that patients to get to a higher functional level have to have some predictability, a stable platform, before they can get to the quote "next level." And that happens in arthroplasty. So I would give you that as sort of a natural experiment with data that says that you can disentangle, even though their pain is sort of the same, they can achieve a better functional level on standardized measures.

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: Just to continue the friendly debate. I would see it the other way and say that most of

the studies are enrolling people who are taken off drug and allowed to flare so pain is a major component of entering into the current OA studies. There are other kinds of patients that we're now thinking about treating that don't come to us with a lot of pain. The one that comes to mind as an example is a woman I saw who is 50 years old the other day who had early OA, doesn't have a lot of pain, but became concerned with a little bit of pain that she had, couldn't walk, some of her functional disability. And what I wanted to give her was not a drug that would take away her pain so much, but I wanted to give her a drug--did we have a chondro-protective kind of drug for OA where what I wanted to measure was the rate at which she would be expected to deteriorate over the next one to three years. And so therefore pain was not the reason I wanted to get an OA drug for her. It was really function, I guess function on my prediction that she would have structural deterioration. I don't know if that is the right kind of response, but it's a different class of patients that we now are looking for drugs to treat with OA.

CHAIRPERSON PETRI: Dr. White.

DR. WHITE: I have a question for Kent. If you could explain what you would view as meaningful in terms of functional improvement because somehow that judgment had to

come in in how you decided six months because I would say would you not view it as important or useful, beneficial, if a patient where we have pain and function linked, if a patient took an agent that did improve pain and because of pain did have an improvement in function and that functional improvement was significant at one month or two months? Would that not be adequate to give you a functional claim for, say, short term? How would you deal with that?

DR. JOHNSON: Well, what entails a meaningful improvement is actually another big issue that we haven't touched on, and, you know, is two points change in your blood pressure significant? You know is five points in your cholesterol significant? Is two joints out of 20 joints in a rheumatoid significant? That's a different issue, and you can get around that by driving it with huge trial sizes. So that issue aside, we're simply looking at improvement as shown statistically in any of the validated functional measures or any of the validated pain measures, and, you know, in essence, if you've got a drug that these things are linked, then you end up getting both claims would be what would happen.

But the question comes up, and it really is a question with ramifications more in longer-term drugs like the chondro-protectors and all that where you may more

dissociate pain from function, and it gets back to this issue of to what degree do we co-require the other variable to at least not deteriorate. But does that answer what you were--

DR. WHITE: Well, sort, but again I would say that by putting that duration, you may be making it quite difficult? Are you making it difficult for agents such as nonsteroidals to get a claim for functional improvement?

DR. JOHNSON: Okay. I'm sorry. Yeah. Yeah. The other dimension of this is the arbitrariness of it all, and we admit that what's in there is arbitrary. In fact, it was just out there for discussion. If you look at nonsteroidal trials--somebody was commenting on this before--you do get most of the effect in a week, but it takes about a month to get the full effect. At least that's been my experience from looking at these nonsteroidal trials. So to push a month to six weeks is not a big deal. The issue comes up, the issue really is your comfort with labeling a drug if you've got the ICH experience, let's say the minimal ICH experience, whatever it is, for six months and a year, and if you've got two three-month studies, is that enough comfort or do we want, you know, if you're going after a function claim, is that enough comfort or do we want something that is a little more rigorous that goes out

because of the things that Nigel was talking about? Is there enough of a concern that we want more rigorous data, not just open data, at six months?

DR. SCHWIETERMAN: Let me just add to that. I think that Kent raises, and this is a sticky issue, and I think you raise a good point. We're not disputing the notion that a month's worth of functional improvement is worthwhile because if I'm better for a month where I wouldn't have been, I'm better off. The question is can you discriminate to a degree between the experimental and the control so that you're satisfied that what you've seen after only a month's worth of data that you have a drug that can be given out to the general public? But our experience has been that it's very difficult sometimes with these subjective endpoints to interpret the data after only a short amount of time. Obviously, it depends on the drug, and if there are new classes of drugs, this may all become moot. But so far it's not that easy.

CHAIRPERSON PETRI: Is there anyone on the committee that feels strongly that six months is necessary? So I think we're at least heading somewhere. Now we would like to take a break. So we're going to take a 15 minute break and then reconvene.

[Whereupon, a short break was taken.]

CHAIRPERSON PETRI: Thank you. Now, I believe some people still want to make comments about pain and function linkage before we try to move on to structure, and I might actually ask Dr. Schwartz who was just chatting with me if he would like to make his comments public.

DR. SCHWARTZ: Thank you. I'm Ben Schwartz from Searle Clinical Research. I think the point that I'd like to make is that I think the signs and symptoms indication that we have had heretofore has really served us quite well, and I think that for any individual patient, there is a balance between how much pain and how much functional improvement is important to a given individual, and that would be very individualized according to the patient. I think we've heard from several people this morning already that pain and function generally tend to go together. I think they are very difficult to separate out. Even in the guidances that were issued, to have an improvement in pain, you also have to have at least no deterioration in function and hopefully secondary improvement and vice versa.

I just can't really see the reason to split those out at this point in time. In reference to what Steve Abramson said regarding chondro-protection, I agree with that wholeheartedly, but I think that's really a structural issue and not really a functional issue. So I would

actually advocate that we kind of keep the signs and symptoms and not try to--I mean obviously get all the data for pain and function and the domains for that but not to separate it out for separate indications.

CHAIRPERSON PETRI: Thank you. Now I think our problem is that these things are all collinear. They do in our minds, though, represent different domains, and so we can conceive that in the future there might be drugs aimed more at one domain than another. And I think that has been sort of where the conversation has ended but let me again open it up for the committee. People who have strong opinions about keeping these separate or strong opinions about trying to combine them? Let me start with Dr. Liang.

DR. LIANG: Are we voting?

CHAIRPERSON PETRI: No, we are not voting. We are discussing.

DR. LIANG: Well, actually I'm confused. Are we talking about a template for measurement or are we talking about an analytical requirement that you're going to make?

CHAIRPERSON PETRI: Neither. We're talking about a regulatory claim.

DR. LIANG: A regulatory claim. So what is that? Is that a hybrid?

[Laughter.]

DR. LIANG: I mean I don't think you can go out on the street and say you're not dealing with pain and function when you're dealing with a musculoskeletal problem. It violates intuitive and patient sense. Now whether you can do it cleanly and rigorously from either a physiologic or cognitive psychology view, I think is a jump ball, but I think you can't have a thing and not say you're not measuring it, no. See I think it would be different if you were forcing us to say you have to have a combined thing or one or the other.

DR. JOHNSON: It's a labeling claim. I mean, you know, my drug reduces osteoarthritis pain. I mean should that be an advertisement?

DR. LIANG: Well, I think people want to hear that.

CHAIRPERSON PETRI: Well, in fact, don't you think patients are going to want to know that before they want to know about function?

DR. LIANG: Well, I was telling Leigh at the break we had a Saudi princess come to the Brigham for a knee replacement. Her primary goal for putting her body on the table was as a princess she could not use a stick. I go back to the stick. And that's why she had the total and she actually was not very happy with the result because she

still had a limp and needed a stick. We can talk about this for a long time, but I just think that, you know, I think it just really--I don't know what this claim game is actually.

DR. SCHWIETERMAN: That's a valid question. Generally, we ask for--I don't think we're asking you to combine function and pain into an index that is a composite of the two. Rather we're asking whether an endpoint that measures function alone has a winner could be enough to allow for approval of an agent even if there were nothing else supporting that as in no improvement in pain?

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: I mean I think we're going through a lot of theoreticals about pain and function and looking for anecdotes, and the advanced patient who might get pain relief without significant function relief was the kind of person that might break the concept that one always follows the other. As Matt was saying, I guess the implications I was uncertain about as we separate the theoretical discussion is the impact on the labeling, which, you know, I think is a separate discussion. I personally don't fully understand. If the labeling said good for pain and arthritis, are we by this discussion saying by unlinking the alternative saying it has to be for pain and function? What are the real implications, let's say, to labeling and the

impact on the drug? You're saying if it's only for function, well, then it's probably not going to be very successful drug given the current models, and I wouldn't want to suggest that we do anything that disrupts the current labeling structure simply based on this theoretical argument that we're having between differentiating pain and function.

DR. JOHNSON: Let me just make one quick comment. I guess traditionally we've labeled these things for signs and symptoms, which is more the way the Europeans are doing it. In the spirit of trying to be both industry and patient-friendly, we have separated these. And also for the spirit of discussion, which we've engendered a lot here. It was brought up by a couple of people at the break to me that if there is necessarily in certain databases anyway strong correlations between pain and function, it may be a moot point because if you win by one, you're going to win by the other, and evidently there may already be some WOMAC and Lequesne data saying that you can't win by the overall measure unless you've at least won by one and not deteriorated by the other, which is exactly what we want.

So if a company measures the whole WOMAC and they win by the whole WOMAC, they've got both claims. Now there is a regulatory decision that we'll have to make about

should we regulatorily separate the two, and maybe we shouldn't. Maybe we should just go back to symptoms and leave it at that.

CHAIRPERSON PETRI: Dr. Moreland.

DR. MORELAND: I would just like to echo I think the game you're playing is to make it pain versus function is really not going to change my mind as a practicing physician, and I think--I don't know why it's going to give any pharmaceutical company an advantage. The patients are coming to see me because of pain. I'm going to give them a therapy, and so I don't think there is need from a regulatory reason to put in a functional claim. I think you just can't separate them enough that it's going to give anybody an advantage and won't change how I prescribe the drug that's available today. Maybe ten years when we have another group of agents that changes function like an MMP inhibitor but not pain perhaps.

CHAIRPERSON PETRI: Dr. Madrid.

DR. FERNANDEZ-MADRID: I think I would disagree in a friendly manner. I think the tone of the discussion is set by the initial line in the introduction--"Current drug treatment in osteoarthritis is symptomatic." And I think this applies to the drugs that we have now for osteoarthritis and some of the drugs that have been

developed. It applies to the NSAIDs. But it may not apply to drugs that are currently in the pipeline and will be available in the future. We are looking at bone mineral density as a parameter that if influenced could possibly have some effect on osteoarthritis on the long term. We are looking at the inflammation of something that may be affected by drugs other than the current nonsteroidals.

We are, as Steve indicated, looking at chondro-protective drugs.

So answering your question that if a drug could be shown to influence function and not pain, my answer would be yes, and it may not be available today, but may be available tomorrow or the day after tomorrow. And it is very possible that we may have to treat these patients with a chondro-protective drug or a drug that will influence bone mineral density and some other drug that may influence pain. So this would be my answer to that.

CHAIRPERSON PETRI: Dr. Tilley.

DR. TILLEY: Again, I think we're getting into a lack of precision in our terminology that is getting us confused because one of the concerns that I've heard is that the WOMAC doesn't separate pain and function. Well, that doesn't mean that pain and function aren't separate entities. It means that particular instrument is mixing

them together. And, you know, Kent is correct that then with that one if things are mixed up together, if you win on one, you'll probably win on the other. I think what we've been talking about is something more basic which is if you could define a functional measure like using the stick and not using the stick that somehow separated from pain in its measurement properties, then would we require both, let's say, a significant VAS for pain reduction and a significant change in this functional measure? Would we be happy enough with the functional measure? And I think speaking now from the patient perspective, what I would like is I would be happy to know the answer to both questions and if the drug was being marketed because it was improving function, it would also be useful for me to know that there was no improvement detected in pain, but I wouldn't necessarily depending on my lifestyle maybe I wouldn't care about that, but I think part of it is being sure that there is sufficient availability of information for an informed decision by the patient--but trying to separate in our minds what we're really talking about here.

CHAIRPERSON PETRI: Other comments from the audience? Okay. I think we're ready to move on to the next claim which is structure. The primary efficacy variable is currently a comparison of baseline and final radiographic

scores for knee or hip, provided some pain or function improvement is also demonstrated. Trial duration should normally be at least one year. So I think the first thing that we can discuss is whether we agree that X-ray should be the basis of this claim. Dr. Moreland, can I start with you?

DR. MORELAND: I think at this time point that's the only one that has been clearly validated. The arthroscopy and MRI measures, at least to my view, are not there yet so we're left with radiographic changes, but then those need to be detailed radiographic measures as currently the doxycycline study, I think as Matt alluded to, that study with many of the designs and many of the techniques used is going to be pivotal. They're using a trial design with obese women who have knee OA and we're looking at the other knee. That's not an easy study to recruit for. Perhaps we would like to have a better trial design that is more applicable to patients who don't have mild OA and they are not so obese. So I think if you go with a different trial design, then maybe one year is not going to be enough. So I think there are several issues, depending on the patient population you choose to measure and also the technique you use and the sensitivity of that technique. It's very difficult to have reproducibility with that.

CHAIRPERSON PETRI: Let me ask a general question. Is one year sufficient?

DR. MORELAND: So far it's sufficient because that's sort of where we're starting. I think we will be looking at two-year data also in this particular trial to see whether there's a difference between one and two years, assuming we have the drug that's going to make those changes. I mean the question remains whether doxycycline will have those clinical benefits.

DR. JOHNSON: Maybe somebody from the audience has it on their fingertips, as it were, the joint space narrowing data and what it implies about sample size and trial duration because that's the other setting that there is some data anyway.

DR. DOUGADOS: I would like to come back to the problem of the choice of which tool. At this time we have two very dated tools in terms of reproducibility, clinical relevance and sensitivity to change, which are plain X-rays, standardized plain X-rays of the hip and the knee, and arthroscopy. Arthroscopy is much more aggressive than plain X-rays. That is the reason why we have said that we cannot propose arthroscopy as the main outcome measure in the development of a drug, but arthroscopy may be advantageous at several points. In terms of X-ray, I am speaking about

not plain X-ray but standardized plain X-rays with precise recommendation concerning the patient positioning, direction of the beam and training session for the radiologist unit and training session for the observers. After that, we have some experience considering the change over time. At least I have experience with knee and hip osteoarthritis of two years and three years, and usually the most important changes occur within the first year.

The second year and the third year, the change is less important than the first one, and probably this is related because there is a correlation between the changes in the structure and the symptomatic severity of the disease at entry. Usually we focus clinical trials on active patients, activity defined by symptomatic severity, such as pain and functional impairment, and in the long-term epidemiologic studies we have shown that these factors, these clinical factors, pain and functional disability, is of predictive value of structural change. That is more you are painful today, more you will be at risk to progress the next following year.

Okay. But the problem, we must emphasize, we know in this particular subgroup of patients, we know the rate of progression is always .2 millimeter with a standard deviation of .8 or .9, but be careful. In this particular

population of patients, fulfilling the criteria for symptomatic severe disease at entry, if you conduct a population-based study that is in patients without pain or without functional impairment, some of them are painless. Therefore, the rate of progression is very low, .05 or less than .1 millimeter per year for knee osteoarthritis. So the calculation of the sample size has to be taken into account the symptomatic severity at entry so you see there is a correlation between both. But it's possible to calculate the sample size. And based on the experience we have, the recommendation should be even you need to increase the patient, the best is to shorten the duration of the study even if you have to increase. We have a lot of advantage for that. You don't have to follow up--the information after one year is very important. But in terms of clinical endpoint, probably we need two years, at least one year.

CHAIRPERSON PETRI: Before you leave the microphone, may I ask you specifically about MR? Are we so far away from MR as an outcome measure that it's a pie in the sky?

DR. DOUGADOS: We have fantastic slides in the meeting of rheumatology. That is the speakers usually presenting MRI as a potential outcome measure. We have fantastic picture, fantastic slides. The problem at this

time we are still waiting for a scoring system, taking into account the role of amount of cartilage [?] of the disease. And the second point, and we are still missing a longitudinal study. I am aware that in this country, there is a very important longitudinal study on MRI in San Francisco, and we have also primary data. Probably it will be possible to get some very relevant information.

The problem will be the uniformization and the international communication. In other words, the machines the radiologists are using are changing every year or every two years. If you conduct a trial of, a multi-center trial of two years of duration in different countries, that will be very difficult. With plain X-ray, it's much more easier to uniformize the technique. But with MRI, probably in the near future. But we say that for ten years, near future. We are still waiting.

CHAIRPERSON PETRI: Dr. Madrid.

DR. FERNANDEZ-MADRID: I think the X-rays are cheap, it is easy to do, but it is a gradually insensitive instrument, and I could predict that in the next few years, it will be something of the past. And I agree that MRI has been changing. The technique has been perfected, but at the moment we have techniques that are able to measure cartilage with accuracy and reproducibility, and it is possible to do

these. I think it is doable now, now and ten years, and I think MRI is more expensive, but I think it is much more sensitive and it is more likely to correlate with functional or pain parameters.

CHAIRPERSON PETRI: From the audience a comment?

DR. HOLFORD: Nick Holford, Center for Drug Development Science. I would just like to try and bring up an important issue about what it is we're measuring here or what it is I hear people saying. We've heard about the joint space narrowing which is said to be .2 millimeters per year, which can be interpreted to mean it's a slope, which would predict that after five years, we would have a one millimeter change in the joint space. In fact, I think the data on which it's based is really looking at the end of one year and finding that a .2 millimeter change was observed.

What we don't know is whether that slope continues year after year or whether the slope is changing so the shape of the progress of joint narrowing, I believe, is currently unknown, and I think that is what is really what you need to know to evaluate the effect of the drug over time. So I would ask the committee to consider trial designs that examine the rate of progression of joint narrowing if that's the index you're looking at, not simply the change at the end of some specified period of time

because you end up, if you only look at the end of one year, you are completely ignorant up to one year, and you're completely ignorant after one year. You have no data on which you can make any kind of extrapolation or interpolation. So I ask you to consider that issue.

CHAIRPERSON PETRI: In order to do that, is it going to require a more sensitive outcome measure such as MR or arthroscopy?

DR. HOLFORD: No, I don't believe so. I think the point I made earlier is that it needs repeated measures.

CHAIRPERSON PETRI: But to keep patients in a trial for one year is daunting.

DR. HOLFORD: Yes.

CHAIRPERSON PETRI: To keep them in a trial for two or three years may be impossible.

DR. HOLFORD: I understand that, but I would say that if you wish to make a claim at one year that the drug works at one year, then what can you say about using the drug if you use it for two years? The answer is you can claim nothing. So if you only look at one year, then maybe the only thing the FDA should be able to allow claims to say is use the drug for one year. And I don't think that's going to happen. I don't think people will use it that way.

But if you have repeated measures even during the

year so let's say you have three, every three months you make a measurement of joint space narrowing, at least over that year, you will know whether the trajectory is indeed linear or whether, in fact, it shows any nonlinearity suggesting that you are having a flattening of effect and that the effect is not continuing after one year.

CHAIRPERSON PETRI: Comments from the committee in response? Dr. Moreland?

DR. MORELAND: Well, I think I'd echo some of Matt's comments earlier. There is really only one study that's looked at rapid progression, and that's UK study where it was shown that looking in women who were overweight that there was a relative rapid in the contralateral knee. I think most of us would assume that other than that, things happen very slowly. So we've chosen that model to evaluate drugs, and echo your comments about keeping them in longer than a year is tough, especially if they're not coming in because of pain, they're coming in to measure changes, because most of these patients you want to have early OA. They have very few radiographic changes. They will very few symptoms, and so I think the comments that were raised were very important ones and I think there are some that we need to bat around here, but I think the logistics of some of that, we don't know that, and we don't have a sensitive

measure. Obviously you wouldn't want to subject patients to arthroscopy too often if that's your outcome measure. MRI, I think, would be the best, but we don't have the data on the sensitivity of that.

CHAIRPERSON PETRI: Comment from the audience.

DR. BEARY: John Beary, P&G. As we've looked at some of the issues in doing structural trials, just for orientation, I'll resonate with comments I've heard about the value of looking at the biology of each novel class of compounds that might come forth in the next few years to look at structural changes, and I'll invite you to think about the knee joint for orientation, and think about it as a joint organ. It's there as other speakers have said with four millimeters of cartilage. It's narrowing--in data that have been published, at a rate of about 0.2 millimeters per year. You know there's also five centimeters of trabecular bone on either side of the soft stuff that we've got to keep in mind.

And as you look at how you would assess that over longer periods of time, as was mentioned in early OA, the changes are proceeding more slowly so these more lengthy periods of study are called for. As you look at what you can measure it with now, structural change in the knee that is, there is extant information on the use of special knee

radiographs that appear at this point in time to be best validated instrument for measuring the progress of a knee joint as it starts in early disease and over ten or 15 years which won't be the duration of trials, but ends up in joint death, if you will.

As you look at other assessment instruments, as has been mentioned, MRI is still in the process of being looked at, being studied, being validated. There are some challenges to looking at the subchondral bone area, but they're working on addressing this. Bottom line it isn't here right now as you go around and talk to people and talk about how you would conduct a long trial in this disease.

As also was mentioned, cartilage markers are not clearly worked out at this point, too, and presumably they would have to be tied into the structural elements of the joint organ in appropriate ways. Arthroscopy, we've had some commentary on that, and it would certainly have the role more in looking at the cartilage aspect of the joint organ, but be limited in what it could say about the five centimeters or so of bone on either side of the knee joint. So anyway, those are some thoughts that came to mind as I listened to the other speakers address the structural issues, and as I recall you'd already earlier in the meeting dissociated the pain claim from the structure claim. Thank

you.

CHAIRPERSON PETRI: Next comment from the audience?

MR. STEPHENS: Randall Stephens, Putnam Loche. I want to address my comments just for MRI. MRI is not that far away, I think, from being useful in OA. Currently, we're looking at how the cartilage, the bone, and the other structures are affected in OA, and grading systems are being developed. So I believe that not within ten years, as you have said, but within a relatively short period of time, grading systems that can be used longitudinally will be available for research purposes.

CHAIRPERSON PETRI: Before you leave the microphone, can you help us address the issue of duration of a trial for a structure claim if MR is validated, it's reliable in multi-center trials? Would you still want one year for a structure claim?

MR. STEPHENS: I think the shortest duration that you'll be able to see an MR change is probably six months. It depends on where the validation for prediction in MR comes through. If you're looking at changes in signal intensity of the cartilage itself, sometimes you can see that as early as three months. However, being able to determine actual defects may take more like six months.

This is, of course, very early on, but that's probably the earliest that you would be able to say with MR that you're having a change from baseline.

CHAIRPERSON PETRI: Thank you. Next comment from the audience, please.

DR. LEFF: Yes. I'm Richard Leff with Bayer. I had a couple questions that I think are pertinent to the discussion with regard to a structure modifying agent. If I understand correctly, a primary endpoint analysis then would be on joint space narrowing and clinical measures would be secondary endpoints? And--

CHAIRPERSON PETRI: We haven't actually, I think, completely discussed linkage here. We're still stuck on actually how to define a structure claim.

DR. LEFF: Okay. The only longitudinal data I know is on radiographic measurements. And the longitudinal data on MRI, although it's coming available, isn't nearly at the level of knowing the reproducibility in a variety of different patient populations at any given point in time and in the future. From my knowledge of just other people's work on radiographs, it takes several hundred patients to do a two or three year study looking at joint space narrowing as a primary endpoint. Maybe a few less if you take up a select population in hip subjects as I think Dr. Dougados

was referring to, but if you take relatively unselected patients, it takes several hundred patients for two or three years.

CHAIRPERSON PETRI: I think that since it was brought up, we should go ahead and discuss whether we think the structure claim should be linked to anything else, specifically a pain and function claim as well. Dr. Liang, would you like to start?

DR. LIANG: No comment.

CHAIRPERSON PETRI: Well, I'll go to Dr. Madrid.

DR. FERNANDEZ-MADRID: No comment.

CHAIRPERSON PETRI: I can't believe all this silence here. All right. Dr. Moreland.

DR. MORELAND: I guess I have to say something. I think my view at this point based on what we know is that we probably will be looking at a structure claim based on joint space narrowing, and I wouldn't tie that with anything else. If you can show that, we will believe that that's going to alter the long-term course. Again, the comment gets in is one year enough? Do you need two years? Do you need three years? But for simplicity of getting an answer, I think, from a clinical trial standpoint, one year is all you can do from a study, and if you get that, I would keep that alone and not tie it into pain or not tie it into function, but

just tie it into structure. So I'll throw that out for people to comment. Matt is ready now.

DR. LIANG: Well, I just want to go back to my thing--we shouldn't put down the time. I mean if the company came out and said it adds a millimeter over a month, I mean I'd let them report anything as long as you could demonstrate objective data that there was structural, you know, advantage. I mean why a year? This is not something that was ordained, preordained.

CHAIRPERSON PETRI: Okay. Dr. White.

DR. WHITE: It seems to me without being facetious nobody is ever going to go after this claim anyway. I mean it looks to me as though it's daunting. I mean at least right now it takes a long time. It takes a number of patients. The techniques are not good--hundreds of patients over at least a year period of time, and so I would think that there shouldn't be a time restriction because, again, if it can be done by what techniques, and techniques may improve so that it won't take a year, so why put that on? There may be better techniques coming down the line, and similarly if you put a requirement for pain or function, that's going to make it even more difficult to do these studies.

CHAIRPERSON PETRI: My hope is that this is our

new generation of prevention chondro-protective drugs. We want to encourage their development. So, yes, we don't want to make this claim impossible. Dr. Abramson.

DR. ABRAMSON: I agree with what's been said. I think the pain and function will, as in our prior discussion, probably follow along if you improve structure. But, again, this is a separate question that has to be analyzed independent from that. I don't know about the time. I have some worry about making it whenever it can be because I don't know how comparative some of these studies will ultimately need to be, and I'm a little more comfortable with one year.

DR. LIANG: But couldn't you treat that with other language?

DR. JOHNSON: I mean theoretically you could--somebody was saying 700 patients for two years. Well, maybe 1400 patients will do it one year, and maybe 2800 patients will do it in six months. So I mean I think in a sense you're correct that to specify a time is a little artificial unless you've got these concerns about sort of broader issues of risk-benefit that you want to kind of sneak into the trial design.

DR. LIANG: I think the comparison issue is only going to be if another company does a me-too, and at that

point they are going to have to deal with something that has a three month or a six month, one year, and I think it gives incentive for the companies that get to market with something that does something at any time and then follow those patients, and every year they get another gold star, another, you know, they can say that they've got a two-year window now and a three year.

DR. JOHNSON: The problem is going to be, Matt, that there is going to have to be an assessment of the epidemiology that's out there, and people are doing these long-term studies right now to associate MRI with X-ray, and there is going to be an association. And it may not be a great one. So if subsequent to that, a company could do it in a month's time by MRI, is that adequate?

DR. LIANG: Oh, gee, I'd be really fascinated.

DR. JOHNSON: You like that one?

DR. LIANG: I'd buy stock. Wouldn't you? You show that by MR that you decrease fibrillation or whatever, sure. I think the rest might follow, but it would certainly be a--

DR. JOHNSON: Well, that's the issue. Okay. The overlying issue in this whole thing is what does follow and you know the question of how comfortable are we with this surrogate? You know the blood pressure story. There was

tons of epidemiologic data and there were a few interventional trials and subsequent to that, you get your drug approved on the basis of blood pressure changes. In some ways, X-rays are, you know, a more clinically convincing surrogate than blood pressure in my mind anyway. I mean you could sort of, a normal knee is usually normal radiographically, not always. But let's say a five year period where your knee is clinically normal, it's probably going to be X-ray normal for most of that time.

And at the end of 30 years of bad disease, you're going to have a terrible X-ray and you're going to have a terrible clinical disease. So there has got to be some kind of association, but if it's so loose in between that in essence, you know, the unexpected toxicities of the drug are present in an insidious way, then we may find out after three years that we don't have a drug. In other words, that it was an incorrect risk-benefit. Now, to extrapolate, assuming that the surrogate was valid, obviously it's hard to do long-term trials. They do get done. You know in osteoporosis that has been two or three year trials. It's not impossible to do them. It's a question of how comfortable--you know--and the epidemiology is going to keep changing in the next three or four years, too.

And as the epidemiologic evidence becomes

stronger, if it does, then presumably any kind of clinical requirement at the time should become less. We got around this in rheumatoid, as you know, because we sort of used the accelerated approval business whether if you have something dramatic structurally in rheumatoid, your drug gets approved, and you phase four affirm that it does something clinically, and if it doesn't, it can get pulled off using the accelerated approval statutes. But I don't think we can do that in OA. So the issue is should something purely on the basis of structure be approved with no clinical correlates.

DR. LIANG: I think you have to remember that if the company is reaching for the golden ring and wants to do MR and, you know, we now have a quick MR for the spine as cheap as a plain X-ray of the spine. I mean I think the technology is in hand. I don't think that I would require blinding. I mean I might blind the MR guy, but I think that you're not talking about a comparison trial anymore.

DR. JOHNSON: No, well, yeah, you may be right.

DR. LIANG: Sample size and all the other considerations.

DR. JOHNSON: The diabetic retinopathy trials we're talking about, you know, often weren't blind because you got a blinded endpoint. But the issue is control. Do

you have a randomized control that goes on for two years or one year? That's the hard part.

DR. LIANG: Gee whiz. I think that's really putting a lot on.

CHAIRPERSON PETRI: We got lots of comments. I'll start with Dr. Madrid.

DR. FERNANDEZ-MADRID: Well, again, I think relevant to this question is the consideration of the natural history of the disease. And if we are looking at osteoarthritis of the knee, we know that the natural history is a very long history, and from the beginning of the disease that is fuzzy. We don't know really, the patient doesn't know when it started, the physician doesn't know when it starts, but I heard of a series of patients with early osteoarthritis with a ten year average disease, this is not early osteoarthritis. I think the way we will invert the pyramid in the treatment of osteoarthritis, as we did in rheumatoid arthritis, we'll look at this earlier ten years, not at what happens in these hundreds of patients with still normal joint space by conventional X-rays but ten year history of disease. So we will be looking at these earlier ten years where the joint space is still preserved, but there are structural changes that X-rays cannot measure. So we will be looking at other instruments, and I believe that

this is the way that we will do.

So the duration of a study may be completely different if we are looking at these patients in whom probably we can do something more than when we treat patients that have already ten years of disease.

CHAIRPERSON PETRI: Dr. Moreland was next.

DR. MORELAND: I think if we can just step back and look at where we're going to be in ten years from now, let's assume that half of these companies here have a drug that they're going to go, and we're going to have ten MMP inhibitors developed and looked at. In ten years, then how are we going to tell our patients which one to use? So if we come about now and say we'll let MRI with a few changes that haven't been validated at six months get through, we'll let .2 millimeter change over one year make it, which one are we going to give to those patients, assuming that we have ten that make it through the regulatory agency? Which one of those drugs do we tell our patients? So unless we, not necessarily as an FDA community but as a scientific community and a rheumatology community, decide now what the standards are so that in ten years from now when this reality takes place hopefully, we won't--we have to make decisions then and I think they're going to be there. And how are we going to be able to choose to tell your children

which drug they should use at age 40 to prevent them from having development of osteoarthritis? Which one are you going to tell them? So we have to be very careful now with I think very clear data, and I don't know whether the FDA's--I think we need to back up and look at it from a scientific community because that's the reality. And ten years from now which one are we going to be able to tell our patients to use? And if we're comparing apples and oranges, we're not going to know, and I think that's the bottom line.

CHAIRPERSON PETRI: Yes, Dr. Luthra.

DR. LUTHRA: I don't know whether or not there are any biochemical parameters that are being measured to look at generalized OA as an indicator of active disease. Do we know if there is something coming in the horizon which might change the whole way we look at these drugs as well as the outcome of the disease?

DR. LIANG: Many people have tried and failed is what I've heard.

CHAIRPERSON PETRI: Okay. I think it's probably time to move on to the next claim which is durability. I'm going to read this one and then Dr. Johnson is going to translate it.

[Laughter.]

CHAIRPERSON PETRI: The primary efficacy variable

is either pain or function improvement with the other as the secondary variable, along with nonsignal joint patient global measurement, structure improvement, and health-related quality of life assessment. Trial duration should normally be two to five years.

DR. JOHNSON: Who the hell wrote that?

[Laughter.]

DR. JOHNSON: There was one more comment over here, I think, actually about the previous topic.

DR. WITTER: Could I just go back to structure for a second, and just make sure that in terms of primary efficacy variables, is there any discussion in terms of other changes, for example, osteophytes as endpoints? I just want to make sure that we've discussed other types of endpoints you might want to look at.

CHAIRPERSON PETRI: So we had talked about joint space narrowing. Does anyone feel strongly that they want osteophytes as well? Dr. Moreland.

DR. MORELAND: I hope I don't have very many osteophytes, no. I think that's an area I'm not familiar with. Is someone measuring osteophytes as an outcome? I haven't--

DR. LIANG: I think the only data on this is Danielson from Scandinavia, and they've had a long Juneau

cohort, and they had one subset with an osteophyte but no joint space narrowing, and those people were happy at 20 years, and so I don't think it's--but on the other hand, you hear orthopaeds talking about an osteophyte causing pain.

CHAIRPERSON PETRI: That's because they want to take it out.

DR. LIANG: So I actually don't know about--I don't think people have looked at, for instance, subchondral bone cysts and whether it correlates with symptoms better than joint space narrowing, and it really is a kind of an interesting.

DR. JOHNSON: Yeah. Well, Mark Hupper actually was down last week and does have some preliminary data that suggest that osteophytes are a better correlate with function and pain than joint space narrowing, yeah. It's kind of interesting.

DR. LIANG: No one has looked.

DR. JOHNSON: Yeah, well, the risk of arbitrarily--I mean a lot of what we're doing, I'm afraid, is sort of existent methodology driven and what we're missing because, you know, of the tunnel visions of those methodologies is a wide open question, I'm afraid.

CHAIRPERSON PETRI: That sounds as though this is

something that needs to be open-ended. We don't have enough data.

DR. JOHNSON: I think it's definitely worth collecting in trials, you know, the existence of osteophytes, so that we can start to data drive these hunches that people have or don't have.

DR. WITTER: I guess as long as we're here, lest we move and kind of miss at least from my perspective what I'd like to hear, do I hear that what we're trying to encourage sponsors to do then is to collect X-ray as maybe the gold standard and possibly look at other measurements such as MR and markers; is that what I'm hearing in general? That we should be encouraging, especially at the early stage of these kinds of considerations?

CHAIRPERSON PETRI: Well, I would certainly second that. I think a study that doesn't include MR measurements is going to make the field very static.

DR. LIANG: You mean X-ray?

CHAIRPERSON PETRI: No. I think X-ray is a given, but I think we need to develop the MR technology, Matt.

DR. LIANG: I thought your question was whether we were saying everybody should get X-rays as the gold standard irrespective of whether they do arthroscopy or MR?

DR. WITTER: I think part of my question is to

answer how much in terms of what we'd like to see, what we already know meaning looking at X-rays, versus how much do we want to get at some of the questions which I think I'm hearing is we're at such an early phase in understanding these kind of structure/function/pain relationships, should we be encouraging sponsors to look at things that may be more in some minds experimental but in other minds the wave of the future?

CHAIRPERSON PETRI: So I would say a resounding yes, we want to encourage that.

DR. LIANG: Actually I would say a resounding no, because I think--

[Laughter.]

DR. LIANG: No. I mean really if a company could give me an argument that I could analyze that MR is better, I would take it. I mean I don't see why you're hamstringing. If they can demonstrate, I guess the terminology would be if they have a validated measure of structural disease that shows change, I'd buy it. Wouldn't you?

DR. JOHNSON: But they wouldn't have that validated measure unless they had measured it, of course. I mean if it accelerates the development plan and makes it more rational or so on, I would think that it would be in

their interest to do what Jim is suggesting.

DR. LIANG: Yeah. I think a lot of companies are doing that.

DR. JOHNSON: Let's ask some of the companies who are willing to talk to us.

CHAIRPERSON PETRI: Dr. Beary.

DR. BEARY: Well, John Beary, P&G. As we thought about this issue, we thought that the goal of the clinical trial was to demonstrate the safety and efficacy of the molecule. These experimental questions are very interesting, but the only gold standard with longitudinal data are the radiographs. So that is something you can build off now, but in the sense of not having a moving target when you're talking about one year, two year trials, whatever they may be, it's very important to be able to write the protocol and execute it.

The variability problems with MRI at this point have been noted by other speakers, and we agree with that at this point, and technology, and so it would be quite a challenge to incorporate that in a meaningful way into a protocol that was showing, hoping to show structure modification benefits. So I guess point of view I would encourage is to keep in mind what the goal of a clinical trial is and we all should monitor these things, but it may

be NIH investigators and others who first demonstrate any particular new technology be it cartilage markers, be it MRI, be it scintigraphy, be it this or be it that, but there are so many possibilities out there that as you try to be practical in design, have a trial you can execute, this would be a concern if we couldn't go with what they are validated longitudinal data at the present time. Thank you.

CHAIRPERSON PETRI: Next comment from the audience.

DR. DOUGADOS: I would like to come back to the problem of the outcome measures and the usefulness of conducting clinical epidemiological case studies. The first one concerning the comparison between MRI and plain X-ray. At this time all the data--we can give our impression--but all the data we have concerning the main characteristic of these outcome variables in terms of reliability, sensitivity to change, and clinical relevance in terms of correlation existing between the absolute value versus the clinical symptoms or the change during one year versus the change in the clinical symptoms are in favor of the plain X-rays even with the experience of arthroscopy and MRI in some discovering symptoms.

Of course, we can try to improve these techniques, but I would like to disagree with madame, but the plain

X-ray is a good technique and validated technique. The single characteristic which has not been validated but which is very important is the predictive validity. In other words, if we observe a change of .2 millimeter within one year, what does it mean for the future? And this is the key point. This is the reason why within the European Community we have a lot of discussion concerning if we propose a claim which is structure modifying drug, and we agree on that, we agree that if we are able to modify the structure, probably for the future that will be an improvement in clinical condition.

But because of the sensitivity of the new techniques, of the new tools, probably we will be about within one year in a selected group of patients be able to find a statistical significant difference between the placebo and the drug, a statistical significant difference. But what about the clinical relevance? That is the reason why the European Community proposed to accept the claim of structure modifying drug in the case of they must be able to demonstrate that the treatment effect was of clinical relevance. So the problem is the definition of the clinical relevance of the treatment effect, and that is the reason why there is a need for longitudinal epidemiological studies in order to evaluate this predictive validity.

We have an experience in hip osteoarthritis, as an example, if you are able to spare .1 millimeter during one year, therefore, during the two next following years, you will spare five to ten percent of hip replacement. So this kind of longitudinal epidemiological effect are required. In other words, of course, I do agree to increase our knowledge in terms of new tools such as MRI, but I strongly suggest that we only need support in order to conduct long-term epidemiological studies to evaluate the predictive validity, permitting after that to propose the sample size, to propose the range of the treatment effect we are expecting and duration of the study.

DR. JOHNSON: Can I ask you a quick question? This is very interesting. If I understand you correctly, you would be willing to approve a drug with no clinical effects if the change in joint space narrowing met some minimal test; is that right? Some clinically--

DR. DOUGADOS: Yeah.

DR. JOHNSON: --relevant?

DR. DOUGADOS: But don't forget what has been written from the European Agency. It's confusing. Because it's clinically relevant structural effect.

DR. JOHNSON: How do you determine that?

DR. DOUGADOS: If you then conduct first

epidemiological studies in order to evaluate the predictive validity. The predictive validity means you are looking at a change after one or two years, and then you follow your patient during ten years to look at what happens after ten years. Otherwise, it's impossible.

DR. JOHNSON: Okay. So you're going to wait until the epidemiology is robust enough to be able to make a pretty good call about the clinically relevant change in joint space narrowing?

DR. DOUGADOS: Yes, that's the first possibility to answer. The second one is to look at the first preliminary results and to get a consensual approach to get a clinically relevant effect. The other possibility is a consensual approach.

DR. SCHWIETERMAN: Just a point of clarification. Would you support then an accelerated approval type of structure where you would approve a product with questionable clinical relevance? In this particular type of structure, you would demand then that the sponsor conduct clinical trials over a longer term post-marketing? Would that be something that you would support?

DR. DOUGADOS: I don't know whether or not they have to conduct the clinical study in the same group of

patients for when they ask for the registration. But what I say I have not the answer, but I know the problem. That is I am quite convinced that the tools we are using right now will be able within the near future to demonstrate a statistical significant difference between placebo and the drug with a P-value less than .05 if we include selected population of patients, if you include hundreds of patients. The problem will still remain about the clinical relevance.

And it's not appropriate--I tell you the problem. I don't know the answer, and the answer will be consensual meetings, long-term epidemiological studies. Otherwise, if you look carefully at what is written in the European recommendation, it's impossible to answer the question. So I have not the answer.

CHAIRPERSON PETRI: Dr. Liang, you have a last comment here?

DR. LIANG: I think your presentation is confusing, a couple different conceptual domains. And this is the continuum from an impairment at the organ histopathological level before it becomes a symptom, before it affects a patient's function. What we're talking about, when you talk about clinical relevance, I think I'm very comfortable with this. This is my bag, but I think that's downstream in the pathway of causation. When we ask the

patient if you improve "x" points on WOMAC or whatever, was that an important difference? That's what most of us who are doing methodologic research on responsiveness mean as contrasted to sensitivity. But if we had a better measure of the actual physiologic derangement with the impairment side, I think it would be far more sensitive, it would affect sample size requirements, and I would as an article of faith accept that if we do a good job at that impairment end that we will downstream make lives better in terms of pain and function, the clinically relevant side. So I'd like to sort of disaggregate those two because I think when we put them together and put all the requirements of metric metrology on one or the other, I think we're sort of losing the information.

DR. SCHWIETERMAN: Let me just, if I may. The Center for Biologics has put out a document on [?] being manipulated by [?] structural cells that addresses this point, and it's a little bit of a side issue, but this is all getting to be futuristic therapy and some days may be used for things like arthritis. And, in fact that argument was made in this document that with structural endpoints like repairing a bone or something, then you can de facto discern that that is something that is clinically meaningful

to the patient. So I think there is a common ground here. In that same document, however, there is emphasis on the fact that long-term studies are important for these patients so perhaps something similar could be done.

CHAIRPERSON PETRI: If we get back to Dr. Witter's original comment, I don't think Matt and I are that far apart. We would both encourage industry to help develop validated and reliable outcome measures and I think that was what you were bringing up.

DR. JOHNSON: Can I ask Matt an epidemiologic question? Is the blood pressure epidemiology from the past such that you could have done the same thing that Maxime is proposing? In other words, follow patients vis-a-vis their blood pressure over five years and show that those deteriorate by a certain amount have a meaningful increase in vascular events downstream? I mean this is a very intriguing approach. I mean granted it's going to take pretty--

DR. LIANG: If we had a number for rate of cartilage loss, that we could predict?

DR. JOHNSON: Functional loss of joints replacement or something like that?

DR. LIANG: I wouldn't spend a career on it. Because I'm really struck at the other end where we have

people who are already symptomatic, have end-stage disease, that there is such a tremendous variation, and it has to do with the fact that we measure function with a monotonic, with one ruler.

DR. JOHNSON: Okay.

DR. LIANG: But the ruler should be really elastic.

DR. JOHNSON: All right. Well--

DR. LIANG: We're getting to that point, by the way, because we're now doing what people were doing for the standardized tests. I don't know if you've taken one recently, but you can actually take a computer test. Based on the answer to one, select by item theory, item response theory, a question that would challenge you. So you can actually finish the test in an hour and get your final score whereas before you had to answer a lot of batteries. So this is a new error or individualized items, and we're doing that in functional measurement as well. So we can have more sensitive measures that are shorter and that get at, you know, individual functions.

DR. JOHNSON: No. But the function still becomes given that we don't have good validation that this marker, you know, predicts good outcomes or bad outcomes, how do we prevent sort of an incrementalism where a drug gets approved

because they do a trial of 10,000 patients, and they do it over a year's time, and they show there is a .001 difference in the joint space narrowing, but it's statistically significant? I mean what Maxime is saying is he wouldn't approve--I think what he's saying is he wouldn't approve that because it's unlikely that such a small difference is of any clinical relevance.

DR. LIANG: Gee, I could be so glib especially--well--

CHAIRPERSON PETRI: Well, this doesn't surprise us. Is bone mineral density enough or do you have to wait for the fractures? And this is just a general question in new drug development.

DR. LIANG: See I think that we're putting our emphasis on the front end, whereas I think post-market surveillance in the back end is more important. I give some kind of formal blue ribbon for every study done according to your specifications that was one year, two year, five year, and the drug would be, you know, X-1, X-2, X-5.

DR. JOHNSON: Yeah, we're going to do that actually, I think, but we'll talk about that in the next--

DR. LIANG: And I think that would be better, and I think that's the way we could--and then the other, of course, clinical decision is cost and toxicity. I think you

have to provide a level playing field with a reporting and ascertainment of toxicity, but I think you have to--sorry.

DR. SCHWIETERMAN: Just a question. I think that that makes sense. You have a chronic disease. You can't necessarily wait chronic periods before you put something available to the public, but ought you to couple the--as in the accelerated approval regimen, just to get back to that--ought you to couple the requirements to a post-marketing clinical trial, or is it enough to say that you have a .001 difference in a 10,000 patient trial?

DR. LIANG: Well, I've never, is there an example where you approve something and you've been able to enforce the requirement to do post-marketing surveillance?

DR. SCHWIETERMAN: No, this is actually a problem.

DR. JOHNSON: Yeah, no. There is under accelerated approval.

DR. SCHWIETERMAN: Right. Under accelerated approval is the only thing.

DR. LIANG: What's an example?

DR. JOHNSON: Some of the AIDS drugs. One of the AIDS drugs was approved on CD-4 counts with an ongoing NCI trial to affirm it. I don't remember which one it was, but--

DR. SCHWIETERMAN: Right. Beta serum--

CHAIRPERSON PETRI: A comment from the audience.

MR. LIPMAN: Yes. Bruce Lipman from Pfizer. Before we get off the structural claim, I have a couple of questions that I think would be helpful to discuss. One is relating to looking at structural damage instead of as a continuum of millimeters of damage per year or loss of joint space per year, relating it instead to proportions of patients who progress or don't progress. The reason this is of interest to me is we did an analysis a few years ago on published data, joint space with data. That was published in the Journal of Rheumatology. I can't remember the author now. It was several years ago. And in this study, they looked at the average joint space with changes over the course of a year or two, and really there was no significant--it was very difficult for them to get any kind of a P-value that was significant in the number of patients that they looked at.

But if instead you analyze that data by looking at the proportion of patients who had no change versus those who had decreasing width and those who had increasing width, which actually also happened, and then you made some assumptions that you had a drug that was going to influence progression of disease, and do a power calculation, you

could actually do a much smaller study if instead of the endpoint being the average joint space with change, if it was the proportion of patients, if you shifted that distribution of patients with a drug, so that now you had many more in the non-progression category as opposed to the progression category. I'm sure the statisticians here could comment on perhaps why that is, but it was a much smaller number of subjects that you need for that type of a trial.

So one question I'd like the committee and the FDA to address is how would they view data that was given to them in such a fashion where there was a statistically significant difference in the proportion of patients who progressed in terms of X-ray joint space width? And the second question that's not addressed here that I think is addressed somewhat in the European guidelines has to do with the generalizability of data that's derived from subpopulations? So we've talked about the obese woman with one knee and you do the other knee like the doxycycline study. Another one might be sports injury with a cruciate ligament injury and looking at progression of osteoarthritis in that individual or perhaps genetic subsets of subjects that are in the future found to be more predisposed to progress more rapidly? How would people view the generalizability of a claim that you reduced progression of

osteoarthritis if the data was derived from several different subpopulations, let's say, not just one, so as to make it a little more difficult for you?

CHAIRPERSON PETRI: Let's start with your first question and I'll ask our biostatisticians to help. Maybe Dr. Tilley first.

DR. TILLEY: Well, I think he answered his own question. I mean it's really a variability issue. There is a lot of variability in X-ray data and there are times when you can reduce data to a proportion and decrease your variability. So I'm sure that's why your statisticians came up with those answers. Let's turn to the others.

DR. EGGER: It's the only reason that I can see, too, that a categorized response would be more powerful than a continuous response? That there must have been a great deal of variability in that continuous response. I just want to hammer the point that we're all alluding to that statistical significance has to be calibrated with results of clinical importance. Nobody here would use a research tool that was not properly calibrated and statistics are a research tool. And what we're struggling with is what is clinically meaningful in this case? Perhaps epidemiologists can tell us by doing population based studies. Perhaps we can look at the convergent validity of new measures with

measures of joint space narrowing with clinical symptoms and with clinical symptoms long-term or clinical outcomes long-term, but this is a question that we're all struggling with. We don't really know the answer at this point, I think, for this outcome.

CHAIRPERSON PETRI: Dr. Tilley.

DR. TILLEY: Yeah. Also I'm a little bothered by the comparison of this to heart disease because in heart disease we see blood pressure as a surrogate for cardiovascular events and mortality. But here I even hear some confusion about is function a surrogate for structure or is structure a surrogate for function. I mean I don't think we have a definition of that final outcome like we do in cardiovascular disease that makes our life easier in some respects, and that's what--of course, we have to do huge cardiovascular trials to get to that outcome, but at least we know what it is, and I echo my fellow statistician's comment. I don't think, we don't have that measure. What is that ultimate outcome that we're looking for?

CHAIRPERSON PETRI: And unfortunately you've further confused us by telling us that maybe osteophytes are just as important or more important than joint narrowing. So I do think we have a problem. I don't think we want to hold up drug development, the chondro-protective drugs, but

everyone apparently feels quite uneasy and I would share that sense of unease that if it's not joint narrowing that's going to be clinically important, perhaps we better find out what is.

DR. JOHNSON: Can I ask Marlene what she meant? You said that as a matter of necessity you have to at some point make a call about what's clinically important? We never did that in rheumatology, I mean in rheumatoid arthritis, and we approved drugs based on joint counts going from 20 to 19, you know. If the P-value is there, they get approved.

DR. EGGER: I wasn't part of that. In the cooperating clinics, we strive to create measures of clinical importance like meaningful improvement.

DR. JOHNSON: Oh, I know, but the issue is if your test drug, you know, has--if your test drug just barely beats your control, but it's statistically significant because your trial is so large, then we usually don't have any--I'm just speaking for myself--but I usually don't have much of a choice but to declare it a successful trial.

DR. TILLEY: But that's why you have advisory committees, I think, to take into account the clinical and statistical things together. I strongly agree that we should not be just considering the statistical significance

of a result, and I haven't seen that happen in our discussions.

DR. SCHWIETERMAN: And, Kent, I'd make the point it's a lot different when you say that a joint count drops from 20 to 19 than when you said the millimeter of change in the joint space goes from .001 to .002. One is a clinically--it's like a pain scale--you have less pain. You have less pain, you're better off, but if you have a less thin knee than before, it's tough to know how that translates, so I would agree with your comment.

DR. WEINTRAUB: I must say that while there is a some feeling in the FDA that one should take whatever example is provided and say that showing a statistically significant change is present means that the drug will be approved. There are many areas, however, in which the drugs have to establish a clinically meaningful change, and I just want to remind everybody that that's very important. So even though you may get a statistically significant change, if it doesn't reach a level of clinical significance. Now, the level of clinical significance or clinical meaningfulness has to be established before you start the study. That's a critical point, but in any case it doesn't mean that just having a statistically significant change in

any measure will for certain get the drug approved.

CHAIRPERSON PETRI: Dr. Beary.

DR. BEARY: Just to respond to the issue of is the amount of joint space you have relevant clinical endpoint, I think at least in terms of patients in my own clinic who are close to bone on bone or well below the two millimeters of joint space where you start to see a rapid decline to the end-stage condition, and while I'll agree we don't have total body mortality here fortunately, but you could view that end-stage joint disease as joint death. That's the end over that 15-20 year period. And as you look at the end of that natural history, you I think certainly do see clinically the connection of bone on bone, loss of joint space, pain and function. I certainly see it in my clinic.

CHAIRPERSON PETRI: But can I challenge you on that?

DR. BEARY: Yes. You see it in the patients who are doing worse. They come back to you.

DR. BEARY: Granted that's not an epidemiologic observation I've just cited to you, but as I look at epidemiologic literature, I still see that correlation in the severe structural cases of knee OA, probably hip, too. I just don't recall reading those articles recently. But I'm also resonating back to Dr. Abramson's point in the

scenario he raised of when you're seeing somebody early in disease and, as you mentioned, Dr. Madrid, you really don't know when they start with OA as we define it, at some point they do come see you and complain of pain or some symptom, but they probably had the disease awhile. We just don't--a bell doesn't ring when the natural history starts, but his point addressed the issue we got plenty of analgesics, we can take care of the pain now; are we interested in doing something about the structure?

CHAIRPERSON PETRI: Dr. Strand.

DR. STRAND: Thank you. I just wanted to make a quick comment and that is if we harken back to the consensus processes that have been developing, OMERACT and RA, and, in fact, we did through the consensus process define what was a clinically important difference in these different outcome measurements and that was how ultimately we came to have a composite measure. Now we don't have that kind of data yet in OA, which is a big problem. But clearly we also need to have the interest and the incentive to gather the data, and I think you're also trying to develop a guidance document that would support that effort, and so I think, you know, we do want to say that there could be a claim for structure, and that it's going to have to be defined over time as we learn more about what is a clinically meaningful difference.

CHAIRPERSON PETRI: Before you leave the microphone, a structure claim based on a one-year study or a structure claim based on a five-year study?

DR. STRAND: Well, I think that we ought to be a bit practical here. As you pointed out, as an investigator, it's very hard to keep patients in studies for even a year, and so I think that perhaps it's going to have to be something that in the context of what's being learned right now appears to be, I guess, statistically significant for sure but it is viewed as a clinically meaningful measure. Now maybe that is because it is in the context of say signs and symptoms improvement or it's in the context of some other work that's currently going on which will validate what is finally decided as a clinically important difference.

CHAIRPERSON PETRI: Next comment from the audience.

DR. LEFF: Richard Leff again. About a clinically relevant difference, we don't have it in osteoarthritis, and we've had a number of trials for a number of years with short-term symptomatic agents, and so now you're trying to define with a drug you don't even know exists and don't even know what its effects is going to be, what's a clinically relevant difference years from now. To me the analogy is

trying to find out what a clinically relevant HBA-1C is before you had insulin. It's a very difficult sort of task. And so I'd just like to bring that up as a point because even in the short-term trials there's not general agreement on some clinically meaningful difference. In the long-term trials, it would be difficult even in the clinical measures. In the radiographic measures, there is a slowly growing body of evidence that a certain prevention, a certain baseline--excuse me--not prevention but certain baseline value of joint space narrowing can translate into a hard clinical endpoint like the need for joint replacement and the like. But that is not necessarily dependent upon the fact whether your drug will actually produce that effect.

So if you have a drug that stops joint space narrowing by a certain amount, you don't know the quality and what that translates to in the future. I'd just like to bring that up as an issue, because if you do require the five or ten or 15 year joint replacement study and you have no other shorter hurdle to get over before then for drug approval, there probably won't be any candidates to go over that last hurdle.

CHAIRPERSON PETRI: Dr. Schwieterman.

DR. SCHWIETERMAN: Let me just--I've been throwing out these terms. And it struck me that I haven't defined

them. Actually the agency has had to face this problem before with other products, and we've done it both ways. The question is what does this committee want to do? We've had the accelerated approval regs for the AIDs therapies whereby CD-4 cell counts and so forth were viewed as surrogate markers of clinical benefit in the interest of getting these drugs out on to the market sooner. The regulations were adopted for them to say, look, just show that you improve CD-4 cell count and you get approved so long as then you conduct a study phase four and post-marketing risk-benefit. That's accelerated approval. You hadn't demonstrated clinical benefit but you had demonstrated improvement on a surrogate, and we're all aware of the problems associated with surrogates.

We've also said, as I also mentioned, all you have to do with the manipulated [?] structural cell therapies is show improvement on a surrogate like replacing a hole that is in the skin or in the bone because it's so obvious to so many people that you're better off having that hole replaced. That it would be nice if you studied long-term benefits just to see how long it lasts, but we're not going to make you do that necessarily for approval. You get and out and out approval. The question that we're struggling with here is how much weight to put on the surrogate marker

in terms of how much clinical benefit, and the agency has decided that some surrogate markers aren't valid enough to just simply allow the drug to be approved without some evidence while others are. So that's really the issue.

I would wager that we're sort of in the middle here, but probably leaning a little bit toward there are probably is a need for clinical studies long-term if we're using these new validated instruments and if we're getting to joint measurements that are so small that we have concerns about what it's all going to translate to. So I'm just trying to frame it so that you have choices.

CHAIRPERSON PETRI: It's hard to do anything with choices without any knowledge. Another comment from the audience?

MR. HOROWITZ: Zeb Horowitz, Novartis Pharmaceuticals. I don't have answers, but I'd like to raise a couple of other issues with you. I agree with the arguments about using plain films. I agree with the arguments for following markers such as joint space narrowing based on the techniques we have and future techniques, but just as in the bone area with osteoporosis, I would raise the issue about do we have any concerns about cartilage quality and measures of cartilage quality along with imaging techniques, be they arthroscopic or

radiological, for looking just as joint space narrowing?

We may find ourselves in the position a few years from now seeing preservation of joint space narrowing without improvement in function or pain because, in fact, the cartilage that's been preserved is not normal cartilage. We won't know that, and I don't have the answer on how to look at it. I wouldn't want to make it a requirement at this stage that you have, for instance, biopsies. We'd never get patients, but it is something to consider that we don't know how to project into the future qualitative improvements that are physician and patient noticeable until we get a better handle on what the disease process is and how to predict who is going to benefit because even if we can measure a 50 percent reduction in joint space narrowing in one year, two years, three years, et cetera, we have no way of knowing which subset of patients will truly benefit from the therapy.

So we have to be allowed within the confines of drug development to make a claim defined by the experiment that was done and hope that this has benefit long-term over a 15, 20 year period. It's going to be impossible to develop drugs if we have an a priori requirement to demonstrate the final clinical outcome, and I just appeal to you to consider that, but we're not just going to talk about

imaging. I mean we're not just going to have to pay attention to that. We don't have the biological markers yet that are validated. But along as time passes, we're going to have to look at that.

CHAIRPERSON PETRI: Before you can leave the microphone, can you actually address Dr. Schwieterman's choice? Would you feel comfortable with a phase four so there would be accelerated approval with a structure claim, but then there would be a requirement for a phase four?

MR. HOROWITZ: Yeah. I mean representing a pharmaceutical company, my default would be if we have well designed clinical trials with prospective endpoints which are successfully demonstrated in those clinical trials, we would like to be able to claim to match those endpoints. I would not ask for a claim of predicting the risk of joint replacement has been reduced because of reduced joint space narrowing in one year. If I could sell a drug just based on one year of a reduction in joint space narrowing, well, then physicians have judged that there is a benefit in that. I don't think we can go beyond that.

DR. JOHNSON: What about the CAS study? Everybody thought these arrhythmias were, you know--this is the classic example.

MR. HOROWITZ: Yeah, but there is no answer here.

DR. JOHNSON: Well, no, one answer is the proposal that it be phase four validated. That's what Michelle was asking you.

MR. HOROWITZ: Yeah, phase four validation would be reasonable if we knew the best way to validate that. And the only way is with pain and function at this point, but, yes, I mean I think that is the only our company is going to be able to develop drugs. There's too much risk otherwise.

CHAIRPERSON PETRI: I think we need to move on, and Dr. Johnson is still going to have to explain the durability claim.

DR. JOHNSON: Well, this may become not as critical as it seemed to have been in rheumatoid arthritis. Let me just say that that was the spirit behind this. There is going to be a lot of schema that will enable people to legitimately reduce drug development time, and if you're talking about a disease that is 20 or 30 or 40 years in duration, we perceive that it would be nice if there were a target that addressed this therapeutic dimension in patients which is, you know, what am I going to do five years, ten years, 15 years down the pike? So this is just a different kind of hurdle, and as you remember I think a number of the people were part of the rheumatoid deliberations, and there was a pretty strong sentiment that, you know, you can have

X, Y and Z, but you're still not addressing the long-term dimension of rheumatoid arthritis, and this is the same issue here.

So if we think it's important to have this hurdle, period, as one in a hierarchy of claims, how would you define the hurdle? This was our first stab at defining it. I notice that structure is not even mentioned in here. Oh, it is. It is mentioned.

CHAIRPERSON PETRI: Nothing got left out.

DR. JOHNSON: Yeah.

[Laughter.]

DR. JOHNSON: Well, it depends on if you believe in the concept, you just sort of try to figure out how to best define it, and this is what we came up with. We didn't think it would be a valid expression of the concept if you did succeed by pain but your structure went down the tubes. So we did have to have all these secondary non-deterioration requirements which I think is what we put in there. So that was the spirit behind it. And it would work like Matt was wondering, you know. You could have a three-year durability claim or a five-year or a ten-year durability claim, but the minimum we thought should be two years because we hadn't done two years yet. We had just done three, six month and one year.

[Laughter.]

CHAIRPERSON PETRI: Well, let's start to address this. Dr. Madrid.

DR. FERNANDEZ-MADRID: Yeah. I think it is appealing in a disease like osteoarthritis to have a durability claim of at least three years. I would support that and sometimes I think three to five years may not be unreasonable.

CHAIRPERSON PETRI: This would be a way to get from that original three month trial to a longer trial. But I ask how are we going to keep patients in a trial longer than a year?

DR. FERNANDEZ-MADRID: Well,l this is very difficult. It may not be doable for a five year period, but it may be necessary for some, for instance, for side effects, for adverse effects on bone, for instance. This may be very difficult to pick up in short trials. Some of these, I think since you put everything here, I think it is not unreasonable for a three year trial.

CHAIRPERSON PETRI: Dr. White.

DR. WHITE: Just to comment about that, though, from a different point of view and a question for you, Kent. Is there a reason to believe that durability and what would be judged to be meaningful should be the same for pain as

for structural problems? And it would be my first thought that perhaps a durable medically meaningful result for pain might not need to be as long as three years, that that might not have the same requirements.

DR. JOHNSON: Yeah. There is probably no reason to assume that it should be the same. The problem with this is a lot of us sort of in theory or in principle felt that we should offer it as a claim, but once you do that, then you realize how hard it is so we tried to make it as loose as possible. You can have any one you want. You can improve in anything you want to measure, and you just don't have to go down the tubes with regards to the other things. Make it the path of least resistance once you've established this long duration. That was the spirit behind it.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: Small point. I think durability is really a tough gate. Look at OA, you know, pristine OA or OA knee replaced by a joint, you know, an artificial joint, there's a natural trajectory that sort of overlaps normal decline in function. And so I don't like the word "durability" because I think that none of these are going to be durable because, you know, once the process gets set up, especially in the knee, biomechanical factors probably accelerate the damage. I mean I don't know for sure. But I

think that is a likely scenario. It's certainly observed. So I just think that that's, I don't see why you can't just tuck that in into the structural thing and just say or have it as a requirement of labeling, not say durability, durable, but just that we've done the good stuff.

DR. JOHNSON: It could be, yeah. I mean you can have a two-year structure, a two-year function, or two year pain or five year pain.

DR. LIANG: I think the real thing is I want to know that they've done a study that you have defined in terms of measuring both the bad stuff and the good stuff two, five, et cetera. I think companies should have a leg up for doing quality outcome assessment like that. I don't think that these are going to be flat trajectories once you do any intervention because it doesn't appear to be that even in the best interventions we have.

DR. JOHNSON: I think we were hoping that by making it a separate claim, it would offer more attraction to industry.

CHAIRPERSON PETRI: But realistically speaking, how many patients are going to be left in the trial five years down the road?

DR. CALLAHAN: And do you have that tied up--

DR. JOHNSON: It's not quite like in rheumatoid

arthritis, you know, where there are so many different. Everybody is going to have all the background therapy that's around, and you know you'll have an MMP inhibitor which doesn't kick in for a year anyway and maybe it's just a mild kick in after that. I mean I think this does bespeak certain types of drugs and not others, but I don't think in principle two year trials are impossible. I mean they have been done in other fields.

CHAIRPERSON PETRI: We have lots of opinions from the audience. Dr. Strand, first.

DR. STRAND: I think I won't offer an opinion. I'll ask a question. And that is what's the definition of no deterioration because I think that's signal component to this whole point, and the other point is why are we worrying about another joint as opposed to the joint that the patient is either symptomatic with or dysfunctional with or structurally deficient with, in a sense, because we don't necessarily know how many people are going to develop bilateral disease nor are we necessarily talking about systemic therapies.

CHAIRPERSON PETRI: Although it was brought up before the generalizability of what we're talking about, the drug that works for knee OA might not have any effect on Heberden's nodes, for example, so that nonsignal joint

problem is an issue that pervades everything we've talked about. Other comments on Dr. Strand's opinion and question? Dr. Gorelich.

DR. GORELICH: Ken Gorelich from DuPont Merck, and I like conceptually the idea of a durability claim. I think it's something very valid. One of the problems that I cannot see overcoming in achieving that kind of a claim is meeting the regulatory requirements of adequate and well controlled clinical trials to demonstrate that kind of a claim. At the same time, we have a natural history that tells us that this disease progresses and so a way perhaps around this where you can maintain some kind of numbers in study is to avoid, you know, the more traditional controlled study structure and look at an open label population with limited sampling over a period of time, and that way you don't have the negative impact of being on a potential placebo. You're on open label drug and I think those are factors which will enhance patient compliance with the study. The question is would the agency be interested in allowing a claim based on that type of open label uncontrolled long-term data?

CHAIRPERSON PETRI: We call those registries.

DR. WEINTRAUB: Or maybe not open label but we also call them large and simple trials, what you're wanting

to do is perhaps with a very limited look at a particular problem once a year or something like that, large and simple trial.

CHAIRPERSON PETRI: Next comment from the audience.

DR. LEFF: Dr. Leff again from Bayer. With regard to the durability, it's a nice concept. I think as an analogy, because we have longitudinal data on radiographs and people understand the idea of slowing the radiographic progression, and we have an idea that most of the drugs we have improve people, I think that durability is actually demonstration of slowing clinical progression, which we have very little information of with disease specific measures from my understanding. In terms of long-term measurements of WOMAC or Lequesne or other measurements, specifically focusing on osteoarthritis, and we all say and we all admit our patients get worse, but we don't actually have a whole lot of measures of that as compared to what we have in terms of measurements for radiographic progression. And so the durability to me sounds like it's a slowing or halting clinical progression over time which is comparable to that in radiographic progression and slowing that which we have a lot more data on.

CHAIRPERSON PETRI: Dr. Beary.

DR. BEARY: I have a question regarding if somebody could explain the nonsignal joint patient global concept and how that would be implemented and used in a study? Thank you.

DR. JOHNSON: We haven't talked about this and it's an awkward notion, and there are probably better ways to get at it, but the gist is to figure out what's going on with the rest of the patient from an articular point of view. You could say just let that fall out in safety. If some of the MMP inhibitors have the risk at too high of a dose of engendering fibrosis and frozen shoulders or something like that, and you're treating a knee, then it might be important to pick that up. Maybe it was an overly construed notion, but it's just a way to remind us that at least we should collect data that pertains to the rest of the joints, too, in some sense. You know Bellamy addresses this too in his letter, by the way.

We haven't had any--you brought that up once before, Michelle, and we didn't have much discussion on it, but maybe it's a faulty concept and we should reconstrue it, but--

CHAIRPERSON PETRI: I think probably people do have some comments on the nonsignal joint. For example, I think Dave Felson had several comments in his letter,

concern about including nonsignal joints in these claims. My original concern was I have no reason to suppose that knee, hip and Heberden's nodes are the same process. Other comments? Dr. Moreland?

DR. MORELAND: I'd just comment about the durability. I don't see how that's feasible at all as far as a claim and I understand where you're coming from and what you tried to offer, but from a clinical trial standpoint, and from a claim, I don't think it's going to be a useful mechanism.

CHAIRPERSON PETRI: Although using Dr. Liang's terminology, the gold stars, if someone wants to do a one-year trial instead of a three month-trial, that is sort of like a pain plus or a function plus claim.

DR. MORELAND: Well, I think if you're in the realm of a three month versus a year, but if you're talking about a one-year versus a five, I think the registry or however you want to call an open label trial that was alluded to, obviously there are so many inherent biases put into that, many patients go into studies to get free medicines. And so, yes, they're going to be durable and they're going to stay in that study for the free medicine, not because the drug or device was durable, but because it was financially important for them to stay in.

CHAIRPERSON PETRI: Okay. Now I'm going to move on to the next claim, which was delay in new OA development. Survival design should include time-to-event analyses. The agency is asking for comment on whether a duration should be specified and if so what duration is appropriate. And in the additional notes that I got from Dr. Johnson, his first point was whether this was practical? Can one be assured of no new OA anywhere without prohibitively extensive X-rays? And of course, the second point is what should be the trial duration?

DR. LIANG: I move that we nix this.

DR. CALLAHAN: I do, too.

DR. LIANG: This is a really a tar baby.

DR. CALLAHAN: It's impossible.

CHAIRPERSON PETRI: Say more than that.

DR. LIANG: Well, I'm sorry I'll put on my academic hat. We can't date the onset of OA in OA patients anyway; right? We get them probably 20 years into their incubation period when they start to have symptoms, but we know that the histopathology begins before that. So I don't know how you're going to do this.

CHAIRPERSON PETRI: Dr. Callahan.

DR. SCHWIETERMAN: It's the Ken Brandt model.

DR. LIANG: Pardon?

DR. SCHWIETERMAN: It's the Ken Brandt model, you know.

DR. LIANG: The doxy knee?

DR. SCHWIETERMAN: Yeah.

DR. LIANG: You know that's based on 47 patients or something.

DR. JOHNSON: Well, I mean it's a different issue as to whether the robust but is the concept robust. You know if you've got one bad knee and the other knee looks normal clinically and radiographically, it sounds to me like that it's--

DR. LIANG: Yeah, but, see, sometimes what happens is you make one knee pain free and that means they're more active, and, they're, you know--

DR. JOHNSON: Well, that's why you have a control.

DR. LIANG: Well, I'm not sure. This is the same body though, you see, with different biomechanics. This is not really a control, I don't think.

CHAIRPERSON PETRI: Dr. Callahan.

DR. CALLAHAN: I disagree. I think it would be very difficult to do that based on--I mean you don't know anything about the rate or the differences in people and there would be so many other confounding variables with the

individuals.

DR. JOHNSON: Well, there are people who feel that there is a data to support these things, and, you know, for everyone that is confounded within the patient, your having a control will presumably help separate out drug effect from other effects, I would have thought.

DR. LIANG: Well, like if your knee gets better and you get another Heberden's nodes, I mean I think that's a pretty good tradeoff, you know what I mean? It's just, I think it's going to be really difficult. I mean I guess you could do it. You could try, but--

DR. JOHNSON: It strikes me as more clinically attractive than joint space narrowing, to tell you the truth. It strikes me as the equivalent in OA of prevention of erosions in a hand that's radiographically normal at the outset in RA, but maybe some people from industry have some thoughts on this one way or the other.

CHAIRPERSON PETRI: Dr. Moreland has a comment.

DR. MORELAND: Being an investigator in the current doxycycline study, the logistics of this cannot be put under the table. Patients have maybe the onset of OA in one knee a few months earlier, but catching them at that right time and then looking at the radiographs, whether it's mild versus a little bit worse than mild, it's logistically

hard to recruit patients that meet these specified criteria. The concept, I think, is good, and we'll get the study done and we'll find those patients, but it's not as simple as it looks in paper. Patients tend to have two knees and they tend to both have OA at the same time. And so picking the one out just at the right time that the other one is not hurting is fraught with all kinds of problems.

CHAIRPERSON PETRI: Dr. Egger.

DR. FERNANDEZ-MADRID: I agree that this would be very difficult, but I think the study of the British on 500 cases showed that an index knee osteoarthritis is accompanied with a contralateral structural changes in a very large proportion of patients. However, the evolution of these changes is very slow so I think it would be very difficult to do the study even in the case of the knee with contralateral changes which are already ongoing although not symptomatic.

CHAIRPERSON PETRI: Dr. Egger.

DR. EGGER: I have two points, but I forgot one. My question is if we--it's an interesting concept, and it's one that I have been fascinated with in other diseases. My question is I think it would be very hard to do to operationalize and to be clear about what had been done. Are we going to rule out the possibility that a drug company

can make this claim if we tell you that we think this is very hard and it's not clear how they would do it? So what's the regulatory implication of our advice to you here?

DR. SCHWIETERMAN: Can I answer that, Kent? I don't think we would rule that out at all. In fact, I was going to make the point earlier the whole point of this guidance document is to provide incentives to sponsors to develop drugs in the way that we think they ought to be developed. But if a company comes to us with any proposal that's reasonable and we say go for it, it's just that they need guidance regarding what are the more important things. So the answer to your question is we wouldn't prohibit it.

CHAIRPERSON PETRI: Dr. White.

DR. WHITE: I'd like to ask a question to follow up actually on what Matt brought up, and that is how you would define new OA for this kind of claim? Would you define it radiographically or would you say development of pain or would you say development of functional changes? I mean given that all of these are components, what is new OA that you would use as a definition?

DR. JOHNSON: That would have to be, you'd have to get a consensus on that. I mean if you feel that the concept is valuable, there probably could be, I suspect you could get consensus, but--and obviously you're not going to

have any data heritage to use to power your trial or anything else.

CHAIRPERSON PETRI: Dr. Beary.

DR. BEARY: Our view as we analyze this problem in the context of current OA clinical trial knowledge and making best guesses as to what might be ahead is that the logistics are very formidable, even doing quote "a simple one or two year chronic disease study" takes such a big bite out of your R&D director's resources that the comments about adding phase four and adding this and adding that, I think, throw a little chill over some of us who have to think practically how we execute these things. So I don't want to overanalyze that, but I do appreciate those who are asking questions is it practical, will it work, do logistics make sense, is it affordable? You know 16 percent of the patients are noncompliers in any clinical trial so over the long study with dropouts, all these kinds of factors that we're already grappling with, any new issues do raise some concerns.

In this regard, they're mainly about how would you do it, formidable logistics. As Dr. Moreland has found out in the positioning of the knees in these patients, it takes a lot of training of the X-ray staff. It's not easy to do. It can be done, but once you start adding other issues, you

really start to wonder if you can execute the study. Thank you.

CHAIRPERSON PETRI: Dr. Tilley.

DR. TILLEY: Just a minor point from the statistical point of view. There's an issue of frequency. If you're going to actually use time-to-event, you really have to have observations at some frequency if you wanted to know when the event actually happened. And so I'm not sure even if you were going to move to this what the advantage would be given the kind of think you're looking at over just looking at everybody at some point in time and saying did it happen or didn't it happen.

DR. JOHNSON: You mean you can't X-ray everybody once a month?

DR. TILLEY: Right, right.

DR. JOHNSON: Yeah.

DR. TILLEY: So you really don't know what the time is, you know, except in very gross increments. So that I'm not sure that you need to do this this way. I mean I think you might be able to just ask the question at some fixed point in time.

CHAIRPERSON PETRI: Next comment.

DR. DOUGADOS: Just to report the discussion we had because I am here as a member of the Osteoarthritis

Research Society member, and just to give some comments. The first one was that the concept of new OA is very interesting. Only if you focus the study in patients with osteoarthritis--it is not a population-based study--but in patients with osteoarthritis in a signal joint, you are looking at the occurrence or the new occurrence of new osteoarthritis. There are two possibilities. The first one is based on the study conducted in UK showing that within three years in patients female obese with uni-knee osteoarthritis, there is a 22 percent occurrence of contralateral knee osteoarthritis defined by X-ray. So the conclusion of the discussion was the definition would not be based on clinical and radiological findings but only on radiological findings, and for this, that is the definition of OA, should be based on osteophytes and not joint space narrowing. That was the discussion.

And also another comment concerning the problem of a new OA and the problem of the other joint to be evaluated. Some people have proposed if you conduct a study evaluating, as an example, joint space narrowing in patients with knee osteoarthritis, to systematically take an X-ray in order to evaluate not only the target joint, as to the joint space narrowing of the knee, but also the occurrence of new OA at the end level. And you see what I mean? That is you are

looking at the secondary endpoint the probability of the occurrence of new OA at the end level in patients suffering from knee osteoarthritis. That has been also proposition.

DR. JOHNSON: I'm sorry. The X-ray of what? Of the hand?

DR. DOUGADOS: Of the hand.

DR. JOHNSON: Hand.

DR. DOUGADOS: Permitting during the trial to calculate the percentage of new OA, but not as a primary variable as you propose, as a secondary variable in the structure modifying--that's different.

DR. JOHNSON: So it might be the case that in a couple of years, there's pretty good MRI and joint space narrowing correlation or something like that and you can MRI somebody and just look for new MRI changes if there is a robust association with those with clinical disease. I mean we didn't think through this. We just wanted to throw this concept out to get some response from people. And like Bill says, you know, if somebody comes across with a very credible proposal that, you know, clinically rings true, we'd be hard-pressed to turn it down.

DR. DOUGADOS: This concept is from a personal point of view is very interesting, but again what would be the definition of new OA because based on the study

conducted [?] and I assume the study conducted here, coordinated with Ken Brandt, the definition is the presence of osteophytes, and nobody knows whether or not osteophyte is good or not for the patient. So in terms of end systems, it's very difficult to propose a claim except if you clearly define what is a new osteoarthritis.

CHAIRPERSON PETRI: Okay. We need to move on to the next claim which is delay in surgical joint replacement. Survival design should include time-to-event analyses. The agency is asking for comment on whether a duration should be specified and if so, what duration is appropriate? Let me start with Dr. Madrid.

DR. FERNANDEZ-MADRID: I find this claim extremely interesting, but very, very difficult. I think it is close to not being doable. I think the decision to do a joint replacement like a hip or knee is a very complex decision that may be different in the United States than in Canada or in other medical systems. It depends on the medical system. It depends on the physician. It depends on the patient. And the criteria used in different centers are so different. I find that this would be extremely difficult. For the knee it would be impossible, it seems to me. Maybe in a small subset of patients with hip osteoarthritis with very narrow joint space who are symptomatic already I think you could

predict that surgery is forthcoming soon and it is possible in some subsets but I find this very difficult.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: Two bits of data. I mean ours I think was the fourth largest joint replacement in the world at some point, and when we started applying these quantitative measures of pain and function to the population which I had been taught as a fellow were people who have end-stage joint disease, on these self-reported measures, there's a tremendous variation. The other study is Mary Charleston's where she went and did a population sample using the kind of measures we do preoperatively and she found an equal number of people out there who are not pre-op as the same people who were lined up to get surgery in New York.

So I would say that the data suggests that this is not a discrete endpoint. This is discrete income but not a discrete--

[Laughter.]

DR. JOHNSON: Has anybody done that kind of study in an HMO setting or is anybody thinking about doing that? That--

DR. LIANG: We're doing lots of studies on the subject using administrative data.

DR. JOHNSON: No, I mean in one homogeneous HMO

like Kaiser of California or something like that, one where you could, I don't know--hope that some of these variables would be less prominent. I don't know.

DR. LIANG: That's interesting. I don't know.

CHAIRPERSON PETRI: If there are no comments, the last thing that we wanted to discuss before lunch was the issue of other claims. Any other claims that any of the industry representatives wanted to suggest or that the committee members wanted to suggest? Well, hearing none, we'll adjourn for lunch. We must be back at 1:30, please.

[Whereupon, at 12:15 p.m., the meeting recessed, to reconvene at 1:35 p.m., this same day.]

A F T E R N O O N S E S S I O N

[1:35 p.m.]

CHAIRPERSON PETRI: This afternoon, we're going to review some other sections of the proposed document and we're going to start with a discussion of trial analyses, and I'm going to go ahead and read these paragraphs for those of you who don't have the document.

Certain trial designs mandate certain analyses and may preclude others because a trial in the end is only as persuasive as its analysis. It is important at the design stage to decide what statistical tests are to be done and on what endpoints. Endpoints need to be evaluated by how compelling they are to the clinician and statistical tests assessed by how artificial are the data assumptions they impose. Traditionally OA trial analyses have used statistical tests compared mean changes from baseline and various endpoints with or without adjustments for multiplicity. Alternatively, trial analyses done with a by-patient rating, e.g., better, unchanged, or worse, seem understandable to practitioners. However, by-patient response definitions are difficult to define a priori in protocols because pilot studies are usually inadequate, leaving the risk that post hoc the ratings will prove too skewed one way or another.

So if we could start our discussion with this sort of basic view of whether we should be looking at differences in means or by-patient analyses. Perhaps if I could start with you, Dr. Tilley?

DR. TILLEY: Well, maybe I'm a heretic here, but there has been a tremendous amount of work that has been put into the statistical guidelines for clinical trials that the FDA has written itself already. And my tendency would be unless there are really specific issues to OA to really leave it to that. I mean the issues of to categorize or not to categorize, all of those things are taken care of in power analyses. So that's--but I'll leave it--maybe Kent, you could say a little more because I was having trouble, you know, with what this section added to what's already been done?

DR. JOHNSON: Well, this is the issue that we've touched on a number of times already this morning. Historically it's just been a problem with standard analyses looking at means because companies will measure 15 variables because we can't tell them which are the two or three most important ones even though we claim to have done that in the past. So they measure all of them and sure enough, even by chance alone, some of them are going to come out positive. And there is always this huge contention at the end about

what the study means unless it's a superb drug, and what doing it to the way, you know, a by-patient test or whatever, there may be some information lost there, and I think that that may be serious, but at least it forces all the debate about how the result is going to be interpreted, what's the meaning of the result going to be. It forces it all up front in the discussion of the design of the trial itself rather than leaving it to the end of the day. That's what I prefer, but--

DR. TILLEY: But I think the guidelines talk about multiplicity and adjusting for multiple comparisons and all that sort of thing. I still feel like maybe I'm missing the point here.

DR. WEINTRAUB: Well, it's true that we have the clinical and statistical guidelines for clinical trials, but even all the time we are changing and we have to change. I'll tell you what. I came to the agency I think it's five years, four months and seven days ago--I'm not keeping track, however.

[Laughter.]

DR. WEINTRAUB: Not very long ago. What I came to was almost I felt that I was sitting on a volcano because I had for a long time, just as Kent was saying, had been writing and talking about by-patient analysis. And here we

always used to say there's the FDA and here's all doctors and patients. And the FDA requires this and doctor/patient require that. And when I came here, I just realized that everything was bubbling under the surface like being on top of a volcano about to erupt. And I think to a certain extent that's still true, and now everybody says--I mean there is always the three phases in any large organization: we've never done it that way; we're not going to do it that way; and I'm not going to get out of this corner. Oh, yeah, we have some days when we do that, but it's not a good idea. And then the third step is we do it all the time.

We're already in the do it all the time stage where we've gotten beyond I'm never going to do it. We've gotten beyond sometimes we do it. We're in the stage of frequently doing it. So we have to look again, open up everything again. Wherever we can, though, I believe we should be looking at individual patient responses summed up. Now, they have special problems, and the special problems have got to be dealt with. That's what I think Kent is asking.

DR. TILLEY: But I guess I would really turn it back to the clinical group and say you tell us as statisticians what's the most meaningful outcome for you and perhaps what you're saying is that the by-patient outcome

like they've moved to in rheumatoid arthritis, the ACR-20 or something like that, is the most informative outcome for you as clinicians. Then we'll figure out how to analyze it as statisticians. I mean I think it's really, I'd rather not have the analysis drive the clinical choice of outcomes.

DR. JOHNSON: Well, I think that's what we're saying, too. I think the issue is to have what's clinically dominant about the disease drive the analysis. And what's clinically dominant about osteoarthritis is nothing in particular, I don't think. I mean there is pain and function. And there are these various domains. I mean that's what the Omeract Group did was sort of describe the domains. And the same thing occurred in rheumatoid, but in rheumatoid they're up to seven variables, and I guess you could do designs and analyses and interpretations of analyses using multiplicity adjustments and all that for seven variables.

But that seems to me letting analytic straightforwardness drive, you know, what the clinicians are forced to interpret as opposed to having the clinicians ahead of time say which is really what the rheumatoid people did was what drives an important difference? What's a significant change in rheumatoid arthritis, and, you know, they looked at about 50 different possible algorithms and

optimized it about what separated drug versus placebo. We don't have that data yet in osteoarthritis, but if we did, we might be able to data drive a respondent index, and then we wouldn't quite have as big of a debate. But in the meantime we've got to make some calls on this.

CHAIRPERSON PETRI: Dr. Strand.

DR. STRAND: I think the difference is simply that we don't have the trials and the data that we had in RA. But it took ten years of working in RA to come up with those seven variables, and those seven variables were reached through a consensus process and then it was very possible to define clinically important change in each of these variables.

We have the variables for OA. We decided on them at Omeract and GREES and ORS, and they are reflected in the EMEA document, too. You don't have the data. We don't have the data to decide what's a clinically important difference or improvement in these variables. But they are still available to us and the instruments to measure them are available. I don't think this is so different.

CHAIRPERSON PETRI: Next comment.

DR. DOUGADOS: Just remember within the Osteoarthritis Research Society there is a standing committee for clinical trials, and within this society

everybody agreed that the most important would be to analyze these studies by patient and not by means. And for this purpose, again, it has been admitted that it will be possible to pool the different domains, evaluating the symptoms in whatever the domains, inflammation, pain, functional disability, and for this purpose, that is to propose coming from our society a set of response criteria. It has been admitted that we will try to give this recommendation based on data and not based only on consensual. So just to inform you at this time, there is an ongoing study evaluating the data we have in previously conducted studies in the past by different drug companies, analyzed by the steering committee within the Osteoarthritis Research Society and probably we will be able to give some answers but not before '99. But the objective is to get a set of response criteria such as with the ACR criteria for rheumatoid arthritis that will be the Osteoarthritis Research Society criteria for osteoarthritis only for clinical symptoms without any information concerning both the structure and the [?] only for a set of response criteria for symptoms.

CHAIRPERSON PETRI: If I understand what you're saying, you're going to have by committee consensus the response criteria, and then you're going to go back to

clinical trial data.

DR. DOUGADOS: What we have asked is the drug company to provide us all the placebo groups in clinical trials they have conducted in the past plus [?] groups. And all this data base will be centralized in Stanford, and we will analyze all these databases in order to propose a set of criteria. But that will be based on data but together with a discussion between the clinician involved with the steering committee and the statistician. So that will be both approaches, a clinical approach and also statistical approach.

CHAIRPERSON PETRI: Dr. Schwieterman.

DR. SCHWIETERMAN: We've had a lot of discussion within the agency about this, and I appreciate your comments, Dr. Tilley, because I think that they are germane. Actually, you know, I have a little bit different take now on this. I think that there are limitations to a patient responder index. I think it's obviously more compelling that a certain number of patients meet a categorical outcome, but is it necessarily the best way to design a clinical study losing the sensitivity of the marker as you go? And furthermore, categorical outcomes--and I was discussing this at lunch--are problematic on the other extreme where you have powerful drugs whereby if you set the

threshold very low, you no longer could discriminate between less powerful and more powerful drugs if you simply set this threshold so you have to do other types of analyses. So I think the committee needs to keep in mind that there are pluses and minuses to each approach and rather than simply say we adopt one versus the other, I think we should keep an open mind.

DR. TILLEY: No. I agree completely. I mean I think one of the issues we've seen in rheumatoid arthritis is that as the placebo response rate has increased, the ACR-20 is starting to have less meaning, and so there are definitely are issues of categorizing data, but I'll still turn it back to the clinical people to, you know, lay the groundwork and give us--

DR. SCHWIETERMAN: But Kent's concern is a real one. Multiplicity of a variety of outcome measures is really what's at issue here because one way to solve the multiplicity problem is just to throw them all together and say you had one outcome or not based on these things, and either you make it or you don't. Do you have comments on that? How to deal with that?

DR. TILLEY: Well, I mean one approach that I've been working on and in the process of writing up is the

global test where you have multiple outcomes and you end up with one test statistic for the set of outcomes. I mean that's one approach. However, in a situation where you don't expect all of the outcomes to go in the same direction, like if you expect that you'll see a big improvement in some functional outcomes and the pain scale may go in the opposite direction, you'll lose power actually with that kind of approach, which may be okay because you don't necessarily want to say this is a great drug. If you're insisting that everything go in the same direction, then it's good to lose power when they go in opposite directions.

So I guess what I'm trying to say is there are statistical approaches to deal with multiple outcomes. There's lots of work being done in that area. And so, you know, you give us the outcomes, we'll give you a rationale statistical solution to the problem, and it will impact your sample size. For example, the example here where you end up with a lot of people. You might have if you categorize data have a ceiling effect where going in you have so many people that are categorized as doing well on this outcome, this variable, that when you get to the outcome, there aren't enough people to change, that's going to impact your sample size.

So you pay prices and part of what we do as statisticians is help you understand the price you're going to pay depending on the outcome you choose, and I mean the traditional price with multiple outcomes is if you had seven to divide the alpha by seven, and you know a lot of people don't want to pay that price. So anyway back to you clinicians.

CHAIRPERSON PETRI: Comments.

DR. HOLFORD: Thank you. Nick Holford, Center for Drug Development Science. I just want to bring up the issue of indeed the kind of outcome you choose is important. If you choose a categorical outcome, generally the conclusions you would draw will be of the regulatory variety. That is you can say yes or no if the drug works or not. And then you've run out of anything else to say because it's a very low information analysis you can perform.

On the other hand, if you have a continuous variable, let's say like a pain score or a joint space narrowing, now you will have some opportunity to explore the dose response relationship, the time course of the response to the things I was talking about earlier. So I ask the committee to consider both kinds of analyses should be considered in parallel. They give you different kinds of information. One is information poor but very valuable to

the regulator. The other one is information rich and very valuable to the patient and the clinician, and trials should be designed to meet these twin goals of satisfying the regulator and the patient and the prescriber.

DR. SCHWIETERMAN: I just have to make one. In some cases, though, the responder indices are not valuable to the regulator because we can't, we have to do further analysis because it doesn't adequately characterize the drug. I don't want to be seen as--the regulatory question really isn't that far from what the clinical question is. There are different considerations, but what's worthwhile to the patient matters to us as much as anybody, and I just want to make that small point.

CHAIRPERSON PETRI: Let me again ask Dr. Tilley how she feels about doing both, the comparison of means and a responder index?

DR. TILLEY: Well, again, I think you have to have a primary question for your trial that you're designing the sample size on and this addressing the question of interest. On the other hand, if we think about something as simple as analysis of variance and we do an overall F-test and we reject the null hypothesis, we then want to go in and see, look at the individual peer-wise comparisons and adjust in some way and see what's going on. So again I think there is

information to be gained, for example, in the global test. If one did the global test, you'd still want to look at how the individual outcome measures are doing. It's just that you wouldn't be giving them the same statistical criteria, you know, critical value necessarily.

So I guess I've always been a proponent of learning as much as you can from your data and starting out with a clearly defined question, answering that question, but then going on to learn as much as you can about what's happening. So I would hate to see a trial where they measured the categorical variable and you could never untangle it and get back to its components. I think that would be not a good thing to do.

CHAIRPERSON PETRI: I have a question about the early drug development because we don't necessarily expect the first drugs developed to be strong. In that kind of situation, isn't a comparison of means going to be better?

DR. JOHNSON: Well, you know, I think yeah. I mean if you're excited about your drug and you're looking for something positive to move on with, you know, you'll probably latch on to it. There is all kinds of interpretive problems early on with ascertaining efficacy particularly unless you've got something that's quite dramatic and kicks in very quickly. I don't know if, I mean I think it's

probably true that you lose information when you construct these responder indices so that, as Bill was saying, the most sensitive measure may still be comparisons of means, and if you're going to go forward with any evidence that you interpret as positive, then it would be useful to do that early on.

It's a big problem with some of these slow-acting drugs. We didn't talk about that this morning, but if it takes a one-year trial to do your dose response studies, you know, which if it takes a year for a clinical endpoint to kick in, then you end up sort of blending your pivotal trials with your dose response trials. I mean that's a whole other topic for discussion, but it's something that I think is part of the dynamic that they're dealing with.

CHAIRPERSON PETRI: Dr. Strand, another comment?

DR. STRAND: I don't want to keep overusing the RA example and trying to say that we want to shove OA into the same mold because we don't, but we don't have any dramatic drugs in RA as of right now, and yet we've been able to see with the biologics and a few other projects that a composite responder analysis is a very effective way of looking at response, and if it is a robust analysis, then mean changes across treatment groups in the individual parameters support the composite by-patient responder analysis. So you get to

look at both. You have a single outcome. You don't have to sacrifice your P-value, and we agreed with the seven that were picked for the ACR responder criteria that they were meaningful and clinically important, and we don't ask that all seven of them improve.

I think we need to think about that in the same context with OA. We're looking at pain, we're looking at function, or we're looking at signs and symptoms. We're also looking at structure. We don't yet know how to put all these things together, but I see no reason why not to be able to consider that you can use a by-patient responder analysis and then to make sure that it's robust to look at the mean changes across the treatment groups in the individual components of that analysis.

DR. SCHWIETERMAN: Well, Vivica, the reason you wouldn't do it is because you lose power with that.

DR. STRAND: No, it's a secondary analysis, only to be supportive of primary.

DR. SCHWIETERMAN: Oh, as a secondary analysis, absolutely. I think it's worthwhile.

DR. STRAND: If the secondary analysis doesn't support the composite by-patient analysis, then you have some reason to call that into question even if it's

statistically significant.

DR. SCHWIETERMAN: So which would be secondary? The by-patient or--

DR. STRAND: By-patient would be first and mean changes across treatment groups in the individual components should be supportive.

DR. SCHWIETERMAN: But that's my point. But that's my point. You're using as your primary analysis the weaker, less sensitive endpoint. Why not do it the other way?

DR. STRAND: Most of the time it's not weaker or less sensitive. It really depends on how different the components of that composite analysis are. If we made them just pain and function, as we tried to argue this morning, it's pretty likely that they would vary together and that wouldn't give you much increased sensitivity. If it's more like RA where tender and swollen joints tend to be vary together and the globals tend to vary together, but other components don't vary together, then you actually get more sensitivity. And again we don't have a lot of trial history in OA so it's harder to put it all together right now. And again you're trying to write a document that is going to be

a lot less specific for that reason.

CHAIRPERSON PETRI: Vivica, my concern remains, though, when you have a continuous variable, you turn it into a categorical variable, you're going to lose power. And especially as we think about the development of chondro-protective drugs, I'm assuming that the initial ones that are tested are not going to be the strongest. We may discourage development if we hold them to strict a test.

DR. TILLEY: No. If the continuous variables are not normally distributed and are highly skewed, you actually will have, could get more power by turning them into categorical variables.

DR. STRAND: Which is what's happened with the RA criteria.

DR. TILLEY: Right, right.

DR. STRAND: The other part of it is that if we're starting with new products, we're going to be doing them against placebo as well as what may be available therapies. So they only have to be better than placebo.

DR. JOHNSON: I think you gain power with a composite if the components of the composite are poorly correlated. I think that was Charlie Goldsmith's point in that article he wrote, and I think that's probably true overall with RA because, you know, it may be that the joint

counts are closely correlated. But amongst the seven, there is a lot of discorrelation.

DR. TILLEY: Yeah, speaking from looking at the global test, the lower the correlation among the respective outcomes, the higher the power of the global test. Because it's as if you're looking at different dimensions and each one is giving you more information. If all of the variables measured exactly the same thing and were perfectly correlated, you'd gain nothing.

DR. JOHNSON: But Bill's point was that if there is more correlation, you may not pick up a drug effect as sensitively with the composite as you do by looking at all the components.

DR. TILLEY: Well, again, it depends on the distributions of those components.

DR. JOHNSON: But I think he's worried about missing potentially effective drugs by having some kind of methodologic standard be too high.

DR. SCHWIETERMAN: My concern is that we're confusing two issues. It's one thing for the doctor to have to know what to do about his or her patient in the office that day to decide whether that patient needs some new intervention, and there should be some agreement as to what's meaningful and what's noise. It's a far different

thing to pick up a treatment effect when you have a control that you can compare differences to. I would wager, for example, that if you showed a statistically significant ACR-15 in a 30,000 patient study that that if it was a non-toxic drug would be something that would be worthwhile to patients even though it was below the ACR-20.

Granted, you get to diminishing returns pretty quickly when you start getting to lower amounts, but if the ACR measures clinically beneficial things like joint counts, like swelling and so forth, painful joints, then having less is better than having more, and I don't see a need for a threshold.

CHAIRPERSON PETRI: Comment?

DR. HOLFORD: Nick Holford again. It seems to me that what I've just heard in the last ten minutes is all about power. And power is something that statisticians and others like to talk about, and companies are interested in so they don't waste their money, but as a clinical pharmacologist, I think the point of a trial is to find out what the right dose is and how long you need to treat somebody for. And I think the committee should be discussing those things. If you're going to advocate a particular form of analysis, a particular design, that design and analysis should be able to answer the question

what's the right dose. Once you've dealt with that, then go and talk about power, but can we talk about the primary issues first of all?

DR. SCHWIETERMAN: Well, I think you're right. I think it's important thing to us, but the truth is most of the time we don't know what the right dose is when we approve a drug. It ends up happening in phase four. I think--

[Laughter.]

DR. SCHWIETERMAN: Maybe I was being too glib.

DR. JOHNSON: That's true. I mean fundamentally differentiating an effect from placebo is a big enough of a challenge. You know we have a very hard time differentiating dose responses. I mean that's just a fact of life in rheumatology and I think in most other chronic slowly symptomatic diseases. It's a lot easier in the cardiovascular or the oncology world where interestingly you have these nice solid clear endpoints.

DR. HOLFORD: Well, if I may contribute, that it is possible to evaluate dose response relationships in these disorders. I have published in rheumatoid arthritis looking at an analysis of a disease modifying drug which did exactly

that looking over a one year period. So a longer term trial. So the methodology exists. The methodology that I would use would be mixed effect nonlinear modeling, nonlinear regression, to determine the dose or concentration effect relationship.

So my reason for rushing back to the microphone was to say, well, the reason why we don't know the dose at the end of phase three is because nobody asked the question when they planned the trial. And this advisory committee is in a position to say be sure you ask that question when you design the trial. And if you do that, then you'll get, there's a chance that you'll get the dose out of the trial. If all you say all I want is a P-value at the end of a trial, and that's what you were discussing the last ten minutes, you won't find out the right dose.

DR. SCHWIETERMAN: I think that that point is very well taken, and we spend the better part of our days actually talking to people trying to get dose responses because we do believe it's important. So I would welcome the committee endorsing that. By the way, the P-value is not the be all and end all, although it may seem that at the FDA. If you have a P-value of .05, there is a lot of things you have to show. So I think what this committee tells us is what's the best way to arrange for these trials so that

we have a composite total database.

CHAIRPERSON PETRI: I think that's the key point. A composite total database among all the different studies, and that is what's been very useful about the ACR-20 is that we can now compare different trials from different sponsors and it would be wonderful to have something set up to to be able to do that in OA as well.

DR. SCHWIETERMAN: Right, and I agree with that. The ACR-20 has helped up a great deal.

CHAIRPERSON PETRI: Let me ask Dr. Liang to comment on this issue of by-patient for OA studies.

DR. LIANG: Well, some of the measures that we're talking about are sort of elastic measures, you know, function and pain, and I think they're relative to people's a lot of things. And it makes sense, I think, because patient, and I sort of take the view that clinical trials are helpful, but not in the office, because we try to individualize patients. So I like to see the data. I mean I don't want to get into the dog fight about how you should analyze your study, but I think I'd like to be able to see it patient by patient, and I would expect that you would see what you normally see, and that is the people who are the worse off have the most to gain, but I think that's useful for me to deal in the office situation.

CHAIRPERSON PETRI: Any other comments about by-patient versus means? Dr. Johnson had some other specific questions for us about general design issues for OA trials. One was the issue of co-therapy or background therapy, specifically the adlib use of NSAIDs and analgesics and how to systematically account for this. Let me ask Dr. Callahan because we're going to be talking a lot here about--are people already doing the OTC therapies? How would you like to see this accounted for in the studies of new drugs?

DR. CALLAHAN: I think just recorded and accounted for.

CHAIRPERSON PETRI: I think one question is do we want people to refine the ACR guidelines for hip and knee OA? Would you like to see the control group all on NSAID and capsaicin?

DR. CALLAHAN: The control group all on NSAID and caps?

CHAIRPERSON PETRI: No, should the control group be following ACR guidelines as opposed to being a true placebo group?

DR. CALLAHAN: I think they would have to have that offered.

CHAIRPERSON PETRI: Dr. Moreland, you're looking

as though you have a comment?

DR. MORELAND: I just want to make a comment about the ACR guidelines. I think guidelines are there for us to look at if we want. They are not anything--I think it's important the ACR puts its stamp of what we think is reasonable as opposed to every pharmaceutical company developing their own guidelines. So I don't think patients have to be on nonsteroidals just because the ACR said that's the next step if some of us believe that nonsteroidals shouldn't be used or capsaicin shouldn't be used because of other issues. So I think you got to have the trial as homogenous as possible, and you either say all of them can take those if they want and you worry about the analysis later, or you say none of those are allowed during this study, and try to make it as clean as we can. I think that's going to be so different based on whatever trial you're doing.

For example, if you're doing an anti-inflammatory trial with OA, you probably would want to eliminate some of those. If you're doing structural modification trial where you don't think those other co-therapies will have a difference on the outcome, then you let those come in to however they want to be used by the patient.

CHAIRPERSON PETRI: Then we're going to have

problems, aren't we, when we're going to be comparing these different trials if everyone is going to have a different control group? Now, again it would be so nice if we could compare. Dr. Egger has a comment.

DR. EGGER: I guess I don't have anything new to say. This is an issue that's been around since these sorts of clinical trials have been around. If things change in the background, you can't tell what happened to the drug that is the centerpiece of the study. And the more you can control for it as you've described either by having everybody's background therapy and behavior and assistive devices be the same or excluding people that are too far from the rest of what people generally use as background therapy. Somehow if you are able to control in those ways, that makes a stronger study. If you're not able to control in those ways, if you can at least document what happens so that the sensible reader can evaluate how much of a difference it could have made, that's useful, too.

Epidemiologists are particularly skilled at looking at the direction of bias in a study, in the results of a study. And sometimes there is a lot of insight there. It seems to me that this committee can recommend a lot of care with background therapy and other background issues, but this is a can of worms that is going to be with us as

long as we study these diseases.

CHAIRPERSON PETRI: Let me ask Dr. Liang for his comments on this. Well?

DR. LIANG: You're picking on me, I think. Actually how about a prosperous recommendation that the control should always be on acetaminophen and we shouldn't have any?

CHAIRPERSON PETRI: Well, I don't consider that preposterous. That's the first line of the ACR guidelines.

DR. LIANG: We don't have any placebo trials in OA anymore because patients are hurting, and we want them to be in trials for longer periods of time.

DR. WHITE: Michelle, I have an issue that actually is related both to the statistical stuff and this idea of controls and should it be specified. I have a concern that if the guidelines go too far in terms of specifying exactly how things should be done in terms of data analysis or exactly what should be used as a control that, in fact, you will discourage innovative, creative ways of development of new better methodology and those kinds of things, and so I just ask that, you know, maybe we could hear discussion of that. How far do you want to go?

CHAIRPERSON PETRI: Well, let me have Drs. Weintraub and Johnson respond to that because I don't think

we have to worry about that.

DR. JOHNSON: Well, yeah, I think this whole issue is something that won't go away but also something we're not going to solve and we shouldn't solve. Inherent in the proposition of a randomized clinical trial is very little leeway that actually might change over time regarding withholding care that you know works. You might be able to get away, you know, and you've got to have adequate informed consent and so on and so forth. But, you know, major symptom relieving background therapy is very hard to withhold unless it's quite explicit and there is no sort of long-term consequences and maybe it's a short period of time. Perhaps you can ethically ask somebody to deal with more pain if they happen to have bad luck and get placebo. But I think, in general, the placebo use issue is an ethical call that designers and physicians and patients have to make, and it really has nothing to do with the FDA, as a matter of fact.

CHAIRPERSON PETRI: Well, I think Dr. White's other concern, though, is if someone comes to you with a wonderful stupendous new innovative study design, you're not going to turn them away.

DR. JOHNSON: Of course not. No, it's just got--

DR. WHITE: Well, I'm not so worried about the

stupendous version. I'm sure they won't do that. I'm more worried about the type they're likely to see.

[Laughter.]

CHAIRPERSON PETRI: She didn't mean it the way it sounded.

DR. LIANG: But we took it that way. I actually have a suggestion, and that is that how about this? At whatever point you want to measure an endpoint, I would have you tell the patient to throw away the stick, the splint, the Tylenol and get on a treadmill and then rate his maximum function or pain. I mean I'd like to see sort of a normalization of these measures by eliminating all those crutches, in a sense, and to get them to do a standardized thing which would assess these things much better.

DR. JOHNSON: Yeah. There have been other proposals like this. You know there are a lot of instruments could dream up. People just have to investigate them and document their performance in which case they would replace the WOMAC.

DR. LIANG: Yeah. And I think it would keep people in there because they can use their stick, whatever, during the times that they're not being observed for the endpoint and then during that period of time just assess it, and I think that would really reduce measurement noise and

be more, you know, reliable.

CHAIRPERSON PETRI: There are a few people who haven't commented so let me ask Dr. Madrid how you feel about this adlib use of acetaminophen, NSAIDs, in these trials?

DR. FERNANDEZ-MADRID: Well, I think your question refers--it is almost identical to the previous one, that is on how the placebo should be treated, and I think this poses many problems, I think not only those that Matt mentioned, but basically you may come up with the synovial effusion that should be tapped. The indication could be one of interarticular steroid injection. I think probably both series of patients should be given this if this is indicated. They should not be deprived from this. Nonsteroidals, would you clarify what is your question on the nonsteroidals?

CHAIRPERSON PETRI: Well, simply that patients who are in this placebo group could, in fact, be in an active control group, and that would include the things that we believe are effective in OA and that would include acetaminophen, capsaicin, NSAIDs.

DR. FERNANDEZ-MADRID: I would not object to the use of capsaicin in both groups. But not in only one group. I think a placebo group should be a placebo group.

CHAIRPERSON PETRI: Let me ask Dr. Harris if he has comments?

DR. HARRIS: Well, I can only say that I agree with what Dr. White says that, indeed, we can't really overspecify, you know, what a placebo group should or should not take. If we do so, it becomes much too rigid and you know really doesn't know--it becomes very rigid.

CHAIRPERSON PETRI: Dr. Abramson, do you want to comment?

DR. ABRAMSON: I guess the only thing that comes to my mind is it's very difficult to have these discussions in the hypothetical or the abstract. And you really have to talk in terms of the specific study that you're designing. As Dr. Moreland said, if you're looking at an anti-inflammatory or Cox 2, then clearly there are study designs out there where you limit the access to other anti-inflammatory drugs. I think if you're looking at MMP inhibitors or one of these other things, you then set up a different set of criteria for those drugs. And within the context of the specific studies have some limitations perhaps so that we're all within the same bounds, but as a global question, I think the answer really depends on the outcome of the study that you have in mind.

CHAIRPERSON PETRI: Dr. Beary, you had a comment?

DR. BEARY: Yes, I think the speakers who put the context into what is the primary question you're asking in the clinical trial give us a structure to think about this. If you're looking in the pain domain, that's where your primary question is, and it's a shorter study, six to eight weeks, whatever it might be, the ethics of thinking through placebo use are different than if you have a one or two year study where you've got patients perhaps already on some stable regime of existing therapy that the doctor has worked out over some period of time, which I think it's to Dr. Moreland's individualization issue that, you know, not everybody can wear a size ten hat so you don't make them go take acetaminophen four times a day which is not easy to do. Capsaicin which some can take; some can't.

So as you work with different patients and see what they need and other modalities you use in OA besides pharmaceuticals, you need to individualize that. And so I could see that as being problematic and impacting on patient retention in a chronic study where it's going to be tremendously challenging to keep them in that long anyway. So I think the thoughtfulness with which you're approaching that question is very useful. Thank you.

CHAIRPERSON PETRI: Let me challenge you because I'm concerned about having each study be individualized. I

would like to see the control groups be comparable in these different studies. And I thought Matt's suggestion about acetaminophen might be one that we could reach a consensus on. Let me ask Dr. Johnson for his view about each study having an individual control?

DR. JOHNSON: Well, I think cross-study comparisons are risky. I mean maybe you can standardize all these things regarding co-therapy, but you can't--I suppose it's possible that if your entry criteria are exactly the same, you know, I just think the ACR-20 has given us the impression that we're able to cross-compare trials, but I think that's still a very risky proposition.

DR. SCHWIETERMAN: I think it's worthwhile to have common endpoints for the different studies by which the ruler which you use. But I would agree with Kent. There's ample data in the literature regarding historical controls and the fact that one group that seems and is defined identically to another performs differently in the different clinical trials. So I'm not trying to say that there is no value to it, but there is limited value to similar control groups.

And I frankly am more concerned about other more pressing issues like dropouts and so forth in a chronic study and happen to agree with the comments about some

liberalization of the control regimen to ameliorate those problems, but I think, Michelle, you're right. To the extent that we can keep them relatively the same, we might be able to make some inferences.

CHAIRPERSON PETRI: So I'd like to be able to tell patients is this new drug better than what's already available. That's why I'm not perfectly comfortable with the idea that I'm going to have the new drugs go up against placebo. Dr. Beary?

DR. BEARY: Just in regard to that, I would say with pain different set of issues. Structure there is nothing available. So there a placebo would be an appropriate comparison. So I think customizing the mechanism of action of these various drugs and the biology of whether you're talking about early OA, late OA, middle OA, is called for here. Thanks.

DR. JOHNSON: Yeah. The more you establish drugs that are clearly efficacious, you can't withhold them. I mean maybe five years from now there will be a couple of those drugs on the market and background therapy will have to entail those just like more and more in rheumatoid trials people are on background methotrexate.

CHAIRPERSON PETRI: Dr. Gorelich had a comment.

DR. GORELICH: Yes. Ken Gorelich, DuPont Merck.

Just a couple of comments about background therapy. If you're looking to compare a new drug of a similar class to old drugs, I think there is a very straightforward path to evaluate that. The issue here is how you evaluate novel drugs, novel classes, novel actions. And I think in that setting, the Declaration of Helsinki is very clear in saying that you can't withhold the standard of care from people who participate in clinical trials, and so if the standard of care is, you know, clearly delineated, then you can standardize it. But I think that as Dr. Johnson pointed out, this is a moving target. One of the problems that we're all going to face has already been faced by the companies that are developing drugs for the field of AIDS where from the day you start your trial to the day you finish it, the standard of care has changed three times. And that doesn't do anything to make trials easier, but I think we're all bound by the same ethical, you know, requirements to teach our patients in the best possible way, which creates the scenario where the proper clinical trials should evaluate the novel agent against placebo with both groups getting an appropriate background of standard of care.

That further complicates the ability to evaluate endpoints, but, you know, that's part of an evolving

science. I think it would be very difficult to convince an institutional review board to withhold standard of care from patients in a clinical trial.

CHAIRPERSON PETRI: The next design question that Dr. Johnson asked us to consider was whether NSAID withdrawal designs were appropriate and he mentioned a few issues: problems with return to baseline, non-drug related noise. Are they a clinically unreal construct? And his example is who in real life withdraws one presumed effective drug before starting another? Let me ask Dr. Johnson if he wants to elaborate on this issue?

DR. JOHNSON: This actually was courtesy of Sahar Dewiche [?] who some of you may know. She couldn't attend today, but she is the rheumatologist with the devices side. And I think it's an interesting question. Janet Woodcock has done a lot of reflecting on the excess noise that she perceives is injected into the types of patients that are accrued and their behavior in trial and in the analysis in these withdrawal design trials because everybody flares and that's artificial, and everybody expects to flare, and I guess part of it is that everybody then has the expectation that they're going to get better as you put the new drug on board, on board half the patients.

And Sahar was asking me to ask you to if people

still want to use these designs? And some people argue that they're unethical too. I think Dave has argued that before. But if you want to use these designs and they're attractive because it suddenly gives you some disease activity to measure, should there be an analysis kind of directed specifically toward this kind of design that would be some sort of longitudinal analysis that I don't know if the statisticians have some thoughts about this or maybe the ethical persuasions are trending enough way from these that this is not going to become an issue in the future? They obviously wouldn't be of any good for drugs that take a long time to get on board, but they still will be used, I suspect, for, you know, nonsteroidals, new formulations of nonsteroidals and so on.

CHAIRPERSON PETRI: Dr. Liang had a comment.

DR. LIANG: Yeah. That's one of the bits of evidence that suggested that acetaminophen would be as good as an NSAID. I think it was Paul De Epps study where they yanked people off of NSAID who had been on it for a long time, and they were happy as clams. So I think what I see is--

DR. JOHNSON: Were they blindly yanked off? I mean out of--

DR. LIANG: No. I don't remember whether it was a

randomization. I think it was just simply before and after. These are people who have been getting, you know, the strips every six months, and they just told them to stop it. And I'm wondering it went--it might have been randomly, but I don't remember exactly, but it's not a given that they're going to quote "flare."

DR. JOHNSON: Well, of course, they're not enrolled if they don't flare when these designs are used and maybe some people in the crowd who have done these studies can tell us what fraction successfully passed this enrollment test.

DR. LIANG: Is that immoral for the statisticians. I mean this is you're yanking after randomization. You're--

DR. JOHNSON: No, no, before randomization. You have a screening--

CHAIRPERSON PETRI: This is part of screening.

DR. JOHNSON: --period.

DR. LIANG: This is part of screening.

DR. JOHNSON: And if you don't flare adequately, you don't get into the study. That's usually how it's done.

CHAIRPERSON PETRI: Isn't that taking a very biased subpopulation of OA?

DR. JOHNSON: Yeah, there is always the challenge of generalizability even if you show it in that group, but

that hasn't been perceived as insurmountable. You know we've used nonsteroidal flare designs for most of the traditional nonsteroidal approvals.

CHAIRPERSON PETRI: Dr. Moreland.

DR. MORELAND: I would just like to make a couple comments. I've participated in several trials in which there was a withdrawal with the flare design. I don't like those studies at all because I think the patients and the physicians and the coordinators all know that they have to flare and so the patients artificially flare more often than we would like to admit. And then often patients who don't flare who really would benefit by being in the study don't get in enrolled. So I think my view is that the nonsteroidals that need to be stopped should be stopped for the one or two or three half-lives of that nonsteroidal because you wouldn't want to mix the biological effect of the old nonsteroidal plus your new study drug at the same time point. So I don't think a flare design is reasonable in today's--what we've learned from previous studies.

CHAIRPERSON PETRI: Other comments? The next issue we're going to discuss is called assembling the evidence, and just to read a few of the key points we want to discuss. More than one claim can be pursued in the same trial and claims can be submitted singly or together.

Because the persuasiveness of trials showing a difference is in general much greater than that of equivalence trials, it is highly desirable for a claim to be demonstrated in at least one trial showing superiority of the test agent, compared to placebo, a lower dose of the agent or an active control. If a claim of superiority over a specific drug is sought, it should be substantiated by two adequate and well controlled trials showing superiority which can also be the basis for demonstration of the product's efficacy.

So we have a couple of things to discuss here. One is whether one claim can be pursued in the same trial? And for example, what if one claim is substantiated in one trial but not in the second? Let me start with Dr. Madrid?

DR. FERNANDEZ-MADRID: Well, I would agree that more than one claim can be pursued in the same trial. I think I don't see any objections on this, particularly if we are looking at pain and function separate, for instance.

CHAIRPERSON PETRI: What if function wins in one trial and not in the second?

DR. FERNANDEZ-MADRID: Well, this may be a problem.

[Laughter.]

CHAIRPERSON PETRI: But we all agree it's a problem. Dr. Egger.

DR. EGGER: The question would be what happened? What was the difference in those two trials? Is one of them underpowered? Is one of them different subgroup of patients? What was different? It seems to me in all of these cases, where we may come up with slender evidence for efficacy of a drug that we need to look at what happened? How do we explain what happened?

CHAIRPERSON PETRI: Dr. Moreland, does your feeling? Does a claim have to win in two different trials?

DR. MORELAND: Well, I share some of the same comments that were just given. I think obviously the answer is yes. If you need two and you've clearly defined your primary outcome and if you haven't met them, you probably shouldn't be coming to the agency with that claim.

CHAIRPERSON PETRI: You would create an exception if one of the studies was underpowered?

DR. MORELAND: No, no.

CHAIRPERSON PETRI: Dr. Harris?

DR. HARRIS: My view is that certainly that is stated, you know, I would have no problem there being one or more claims with a trial. I think it will be up to the people designing the trial to decide what they want. I don't know if I've answered that, but that's my view as stated.

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: I'm just surprised that one trial is sufficient. Usually in clinical or basic science, that first trial is just out there to be proven or disproven. And I'm curious in that is it true that one trial is adequate for a claim to be granted as efficacious?

DR. WEINTRAUB: In the setting where the disease is very serious or where we have lots of previous information or there are many categories, but in many cases one trial is sufficient, but not in all cases. And I want to stress that. Not in all cases. Because what we get is everybody coming in with my trial for my cut finger and, you know, it's one trial, three people, et cetera. But what we have to--so the upshot is that we do recognize that one trial may be sufficient for certain things but only for certain types of questions and types of drugs.

DR. ABRAMSON: What kind of criteria do you establish in terms of patient numbers and duration?

DR. WEINTRAUB: Okay. We have a document, a draft document written and it's out there. It depends a little bit. If you have a lot of information, previous information, if you have a solid mechanism of action, things like that, you'll need fewer patients. You'll need fewer patients for a life threatening disease, et cetera. So I

can't answer your question specifically, but we do have a document out there for one study.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: No comment. Oh, am I supposed to vote?

CHAIRPERSON PETRI: Well, you can comment. We're not voting.

DR. LIANG: The more evidence the better.

CHAIRPERSON PETRI: So that sounds as though you're in favor of two trials?

DR. LIANG: I think it takes three to break a tie. So by definition.

CHAIRPERSON PETRI: Dr. Egger.

DR. EGGER: In case anybody misunderstood my previous comments, I think that science is based on replicated experiments. If we have two experiments that disagree, we need to understand why they disagree. If one of them has too little power, there is not an exception that should be made for that. That's a study with too little power and not very much evidence. It may give us suggestions, but it's not a basis for an exception.

CHAIRPERSON PETRI: Let me ask Dr. Johnson if that addresses his concerns?

DR. JOHNSON: Yeah, I think so. The one trial-two

trial debate has lots of dimensions to it. The extreme being if it's a scenario where you could argue that it's unethical to try to replicate it as a second difference trial, then obviously you shouldn't do the second trial. This is boilerplate from the RA document, as many people might recognize, and part of the reason to put it in there was that one of the drives behind all this is to have a hierarchy of claims, but you know and you can go after one claim and then go after another one after your drug is approved for that first claim if you want to.

Or if you think you got a home run, you can go after two or three of them in the same design. I think testing two or three hypotheses in one design is challenging, but I think it can be done. But that's why we put this in here, and we did have this caveat about-- this is more pertinent for RA--but if you're going to go after an equivalence claim to a drug that's already approved for activity in the disease and you want to market it as good as methotrexate or whatever, then you do need two trials. We haven't--you know, that was kind of a decision that was at least made for the guidance for RA.

DR. WHITE: Michelle, can I ask either you or the other members, I know it sounds great to require two trials, but I wonder about again issues when we're talking about the

category of structural changes, for example, and trials and what it would take to put together a trial to address efficacy related to structural changes in OA. Could I have a sense from the other committee members and maybe the audience how likely that it is that two trials, two such trials would be done given that they're going to probably be fairly large, cumbersome, long, extremely expensive trials? And is then the requirement for two really necessary? Is it just a make-up number or is there really some reason to say there has to be two?

DR. HARRIS: In fact, let me say I think that's an excellent point because somehow we may be thinking pain function alone, and not in fact looking at some of the other possible parameters that we've been considering today that might be more difficult to get at.

CHAIRPERSON PETRI: I think when it came to trials about structure, that was really the sticky wicket today in that we didn't think that there was, at least right now, epidemiologic data to tell us what would be clinically relevant in a structure trial. And so, for example, we weren't sure whether joint reduction was the specific thing or whether osteophytes or both. So Dr. Siegel has a comment.

DR. SIEGEL: There is one distinction that might

be helpful here, which is there's a difference between trying to license a drug for the first time versus trying to get an additional claim once a drug is already licensed for that disease. So, for instance, in rheumatoid arthritis, generally we would expect two trials showing that it's effective for rheumatoid arthritis. But once you had that for signs and symptoms, if a sponsor wanted to show that it was effective at delaying radiographic progression, we would not necessarily expect two trials both showing delay in radiographic progression. Once they had already shown that it was effective for signs and symptoms, one clear trial showing delaying radiographic progression would ordinarily be sufficient. That may be relevant here as well. If you already had a claim of decreasing pain and improving function if you were going after an additional claim of delaying radiographic progression, maybe one convincing trial would be sufficient in that case if it were already licensed.

CHAIRPERSON PETRI: Dr. Moreland.

DR. MORELAND: I guess it comes back to the question, though, with an MMP inhibitor, we would assume it would not be licensed already and so would one study be enough?

DR. SCHWIETERMAN: I hesitate to bring this up

because I don't the language, but the FDA Modernization Act addressed the point of the number of trials it took to get a drug approved, and perhaps we could refer back to that for some guidance on this before we write this particular guidance document. In general, there has been less of a discrepancy between the Center for Biologics and Drugs than has been thought of. There is almost always a need for some sort of early phase two study and a confirmatory phase three since it's unlikely that one single large phase three trial, although it has happened, is enough for approval.

And I would argue that perhaps there is a way of incorporating both the Modernization Act and this notion of replicative science, which I think is important, into this guidance document.

CHAIRPERSON PETRI: Dr. Pucino.

DR. PUCINO: One of the things that could be done with the structural changes is if you have a confirmatory study that the second study would not have to do the phase four post-marketing surveillance so that only the first study you'd look for the changes in function and pain and what not.

CHAIRPERSON PETRI: Is there a comment? Please go ahead.

DR. DOUGADOS: Yes. On replication, from a

scientific point of view, we all agree in our group that one study has to be replicated, but in the field of osteoarthritis some of us have proposed that to take advantage of the requirement of the replication of the study to evaluate another localization or other characteristics of patients. Will you agree on that if you take the decision that you will want two studies demonstrating a treatment effect, but, as an example, one study conducting knee osteoarthritis and another one conducting hip osteoarthritis? So you will have two studies. You will have the replication based on the results but with different characteristics of the study patients.

CHAIRPERSON PETRI: That would certainly help to address the issue of generalizability.

DR. DOUGADOS: Yeah.

CHAIRPERSON PETRI: I'm not sure that that's something that we can control though.

DR. JOHNSON: I think that that would be a wonderful submission.

DR. WEINTRAUB: In fact, we frequently ask for that type of submission.

CHAIRPERSON PETRI: Dr. Moreland.

DR. MORELAND: One comment with regards to repeating that second study during post-marketing. If you

have an MMP inhibitor that's efficacious in the first study, how can you ethically then post-marketing enroll patients into a study, placebo controlled study? Everyone is going to get the real thing.

DR. JOHNSON: Yeah. That's exactly right. I mean we're getting ahead of--we didn't think we'd get these discussions going this far. And there is going to be a lot of thought that has to go into describing these things because it's very hard once something is out there, especially if it's already been shown for signs and symptoms, to withhold it. Everybody is going to want the drug.

DR. MORELAND: The question I would have is obviously there are some MMP inhibitors now in patients in trials. What have the companies been told about at the end of that study? What can they expect?

DR. JOHNSON: Well, that's why we're doing these documents.

DR. MORELAND: So we need to decide.

DR. SCHWIETERMAN: Hey, we want your input. Yes.

CHAIRPERSON PETRI: Dr. Beary, you had a comment?

DR. BEARY: Just to respond to the conversation about studying at different sites. And our current thinking is we have looked at our particular program. We are looking

at each joint area as a separate problem and also have found that the imaging of the hip is not quite as far along as the knee at this point as well. So we are viewing those as separate problems, and as one thinks about sizing studies, you can also think about replication coming in a multi-center context as well, as long as the study is sized appropriately for the question you're trying to answer. But it may be operated, the two studies, on quite similar protocols.

CHAIRPERSON PETRI: The next issue we were asked to discuss is overall risk-benefit assessment. And we had talked a little bit this morning about length of trials for the different claims, and I think we agreed that we thought that three months was reasonable both for pain and function. And I believe we agreed that one year was reasonable for structure because the current standard outcome measurement would be X-rays. We didn't think we could shorten that period of time at this point with inadequate information about MR.

But now the issue is more one of safety. And one of the things we were asked to consider was brought up in this document is the last line. If concerns exist, e.g., from the mechanism of action or from prior experience with similar agents, these may need to be addressed in a phase

four program. And let me ask again for discussion on this point. And let me actually start with Dr. Pucino.

DR. PUCINO: I mean you would always look for phase four trials on most of your drugs coming out, particularly a new agent that you don't have information on. So I would endorse phase four trials absolutely.

CHAIRPERSON PETRI: Do you want to give an idea about the length, especially if we're looking at a structure claim?

DR. PUCINO: I think it will depend on the agents again. If we're talking analgesic, then you certainly would have information within six months, within three months even if you're looking at changes in the GI tract. If you're looking at something that may cause fibrosis, that may be different. You may have to look out to at least a year and probably a couple of years.

CHAIRPERSON PETRI: My concern is for the structure claim. I'm assuming that patients will be on the drug for a very long period of time and therefore I agree with you that the safety concerns would require more phase four. Let me ask Dr. Liang for his thoughts.

DR. LIANG: I don't have anything useful to say. I mean I think you really need to follow all patients forever. I mean we would love to have that information. We

don't do it no matter what we talk about in this room, and we're really looking for that rare event that we haven't thought about so there is no way to really, you know, develop a system for that. So I don't have anything smart to say.

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: Yeah. Obviously we want with new drugs to have some follow-up. I don't know how often you can mandate the phase fours and the cost that that would incur, and we have to obviously be very careful in making those kinds of recommendations. I think a one-year study on a structural outcome, I think, probably would be sufficient in my mind in terms of the side effects that might be seen for most of those drugs. I'm curious as to how many, how often phase four studies are performed on drugs?

DR. JOHNSON: I don't think there has been any for OA. This is a flip off of some concerns we've had in RA that may or may not arise in OA, as a matter of fact. But RA, as you know, is a little different. You know if you can show something pretty dramatic in three to six months but you have to worry about lymphomas or opportunistic infections or this and that. It may or may not be the case that we'll get equally worrisome concerns as these OA development programs go along.

If we do, it's a major problem because just following them doesn't help probably for these rare events. You got to know what the background rate is, which gets into you case controlled studies.

DR. SCHWIETERMAN: In general, the answer to your question from at least a general biologics point of view is very often. For the reasons that I said earlier, that we never know the dose at the time we approve it. We ask sponsors very often to look into that, follow the patients for other endpoints and so forth. Usually it's something that they want to do anyway because they have concerns about or interest in a particular area.

CHAIRPERSON PETRI: Let me the rest of the group also for comments? Dr. Harris, your view about this risk-benefit and how long patients should be watched in phase four?

DR. HARRIS: Well, I, as just said, I think that most pharmaceutical companies presumably monitor their drugs anyway after they're on the market, and you know any unknown, any untoward effects eventually come out without this degree of specification in phase four.

CHAIRPERSON PETRI: Dr. White, your comments?

DR. WHITE: I would agree with that comment.

CHAIRPERSON PETRI: Dr. Moreland, do you have anything to add?

DR. MORELAND: No.

CHAIRPERSON PETRI: Dr. Callahan?

DR. CALLAHAN: I just would agree with Matt. I wish all patients were monitored forever. The problem which Kent pointed out is that you don't have the background. I mean we need to have all patients monitored so that when the patients who are in the trials are monitored, you know what the background is to compare to, and that's a real problem.

CHAIRPERSON PETRI: Dr. Madrid, your thoughts?

DR. FERNANDEZ-MADRID: I essentially agree. I think one or two year trials should capture most of the reactions, probably not the rare ones, and I think probably one should look at the bone also.

CHAIRPERSON PETRI: Let me just make sure that Dr. Johnson found that helpful. Does that address your major concerns?

DR. LIANG: You had to get it out of your system though; right?

DR. JOHNSON: That was just to kind of complete the description of assembling the evidence, I guess.

CHAIRPERSON PETRI: I'd like to turn to the questions for discussion. A few of these we have touched on

this morning. So for those we'll spend less time. The first question is sort of very general, overall feeling about the draft guidelines. Is the overall claim structure fundamentally and appropriately construed? Obviously there are a few that we tore down this morning. So I think durability, delay in new OA development and delay in surgical joint development, we were not happy with that, Dr. Johnson, so those went down the drain. But let me ask the group about the others, whether they are happy with them or whether they have constructive criticisms or suggestions? Let me start with Dr. Madrid.

DR. FERNANDEZ-MADRID: Well, I think going back to your first question, I think the guidelines probably after we discuss them today, they will be appropriate as of today, but probably they will not be appropriate very soon from now. It seems to me that this is a document, that it is in transition, it is a working type of document that is trying to unify a policy for many drugs that have very different effects, for many joints which have a different natural history, and it seems to me that in the future probably these guidelines will have to be addressed with more specific group of drugs and to more specific sites. That will by answer to the first one.

CHAIRPERSON PETRI: Dr. Callahan, what was your

view?

DR. CALLAHAN: I pretty much agree with the problems that you highlighted, and the other structure, I think it depends--I mean I assume these were just guidelines. I haven't dealt with this before in that people would deal within these guidelines, but these are not hard and fast bound.

CHAIRPERSON PETRI: Dr. Moreland.

DR. MORELAND: I have nothing to add to that.

CHAIRPERSON PETRI: Dr. White?

DR. WHITE: Again, just my concern that flexibility be maintained because we are trying to hit a moving target.

CHAIRPERSON PETRI: Dr. Harris?

DR. HARRIS: Same. That we need some flexibility. I mean we should be accustomed to that. In the ACR, you know, we have in a number of guidelines that we do, that we deal with, and I think we understand that there is a degree of flexibility, just that we ensure that that be maintained.

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: I've really nothing to add to the discussion.

CHAIRPERSON PETRI: Dr. Liang? Dr. Pucino? Kent, did you have other specific questions about the overall

claim structure? The second question was in OA trials of novel new agents, is it worth trying to capture under the randomized rubric a broader assessment than suggested above? This might be done, for example, by formally defining outcomes described by the patient to include toxicity considerations and just so aimed to have an endpoint closer to the full risk-benefit expression. Kent, I didn't think I really understood this question. Can you explain it to us?

[Laughter.]

DR. JOHNSON: Maybe this is not worth going into. But it would be possible to redefine how we assess patients to capture both efficacy and safety. And then you have randomization as your weapon to propel your inference that bears both on safety and efficacy whereas traditionally we've only used the formal randomized structure for efficacy and let safety fall out. So fundamentally you could make, I mean the thrombolytics have done this in some cases where the salutary endpoints or the deleterious endpoints are all wrapped together in one composite, and at the end of the day you don't even have controversy about the interpretation at trial. You don't even have controversy about approvability of the product because you've done your risk benefit in the trial itself.

CHAIRPERSON PETRI: So your initial sample size

estimate would have to be high enough to include enough of those deleterious outcomes?

DR. JOHNSON: Well, if you had a composite that entailed salutary outcomes and the absence of deleterious outcomes, sure, you're going to have to make a judgment call which is going to very hard because there is going to be no data to drive it. Maybe that's enough to nix the whole notion, but I think--

CHAIRPERSON PETRI: For NSAIDs, I think we have an idea what the deleterious outcomes might be, but for the novel agents such as the chondro-protective agents, I'm not sure we have enough information to do that.

DR. LIANG: But I don't think you should give up. I think that's a really interesting concept, and you could in a sort of structured gamble for the patient--

DR. JOHNSON: Yes.

DR. LIANG: --ask them I'm going to give you Brand X. How much risk of death would you take to get, you know, halfway better and walk to the grocery store? I mean that kind of stuff.

DR. JOHNSON: Yeah. Well, we've done endpoints in other disease settings where there is no heritage also, and you had to have succeeded by A, B and C, and you could not have failed by D, E and F. And if there is enough

experience with the drug that you think you're going to capture most things, and then at the end of the day, all the other passive fallout analyses of safety will not be very meaningful.

DR. LIANG: There is just not enough preference data available. I mean that sort of utility study has been done.

DR. JOHNSON: It's an idea to keep in mind, and it may be appropriate for certain agents that come down the pike in the future. I mean you know it has as much as logical credence, I think, as what we now do, probably more.

CHAIRPERSON PETRI: Dr. Pucino.

DR. PUCINO: Yeah. My concern would be the rare events and how practical that would be for industry to actually conduct the study like that.

DR. JOHNSON: Yeah, it wouldn't pick up rare events. You know you're absolutely right unless you had a giant trial.

CHAIRPERSON PETRI: Dr. Tilley.

DR. TILLEY: I guess the one concern I have when you start combining, you know, positive and negative endpoints is that, you know, are three deaths are they balanced by, you know, five miraculous cures? I mean I don't--

DR. JOHNSON: You're absolutely right.

CHAIRPERSON PETRI: Not in OA.

DR. JOHNSON: And in the thrombolytic world, what they do is if you die or have a heart attack or you have a hemorrhagic stroke, you know, those are all sort of vaguely equally weighted, but I think you're right. That would be one of the many methodologic problems.

CHAIRPERSON PETRI: Then the third question is is there a more elegant way to capture nonsignal joint activity? Given it's strong rationale, should it matter that there is no experience with using such a measure? Kent, let me ask here because this idea of evaluating the nonsignal joints bothers me because I don't necessarily expect a novel drug for knee OA to help Heberden's nodes. Perhaps it won't help hip OA either. Can you give us a little bit more background about how you're thinking about the nonsignal joint?

DR. JOHNSON: Well, again I just want to have that information captured, and I think it's a very awkward way of doing it, and maybe it shouldn't even be a secondary endpoint. But I'd like to at least see it accrued in the process.

CHAIRPERSON PETRI: Does this mean, for example, that you in a knee OA study, you want hip X-rays, hand

X-rays?

DR. JOHNSON: No, I don't think you can justify that. But you can justify asking patients if their hands are feeling worse or if their hips are feeling worse or if you're got a new shoulder fibrosis from your MMP inhibitor or something. I mean it's possible that there are things that would be valuable to capture. They probably shouldn't. As Felson said in his letter, that probably shouldn't have been that dominant of a theme here, but I think it's a conceptual gap in all the ratings from the past even though, you know, when the patients are doing their globals, they're supposed to say globally speaking how does everything bother you that only affects your knee. I mean it's sort of a disconnect in their mind, I think.

CHAIRPERSON PETRI: Comments? Dr. Moreland.

DR. MORELAND: I agree it's going to be a difficult thing at this point to put a handle on it, but I would leave it up to the pharmaceutical companies and others to be innovative in looking at that because there may be some changes in other joints such as the hip or with MMP inhibitors. And so if there are some pilot studies, it should help us advance this field. I think this is--I'm glad you brought it up to help push it along.

CHAIRPERSON PETRI: Well, should it be that

patients should be enrolled in studies who have more than one affected joint? I mean is that one way to address this issue?

DR. JOHNSON: Yeah, I think it is, but as Felson said in his letter, I think that's going to make more heterogeneous. It's probably going to make it harder to show your drug works because the measures get confounded by the presence of symptoms from other joints. Probably the pure single knee OA patient or the pure single hip OA patient is the best one. I don't know.

DR. WHITE: I just wonder, Kent, how would you use the data? Let's say you had a trial and it was focusing on knee OA and it didn't quite reach statistical significance, but the other joints got better, would you then approve it? Or conversely, if you then focused on knee OA and it did terrific but the other stuff got a little bit worse, would you decline it?

DR. JOHNSON: Well, that's exactly right. If we're going to take this concept seriously and if you enroll patients with, you know, all four joints involved and I mean this is inherent in the deciding to use a signal joint approach in the first place, which we have defaulted to because nothing else exists historically. But what you say is always possible, and I think if the data is pretty

persuasive, then we'd have to just say, well, hell, it looks like it works, you know.

CHAIRPERSON PETRI: But let's explore this again because for the structure claim, in fact, I'm quite interested in what's going on in the other joints. If I'm going to commit a patient to the long-term use of a novel agent because I believe the structure claim, I would like to know that it's probably going to help some of the other joints as well. So a structure claim just for knee OA when I know that, you know, a lot of patients with knee OA area going to have other joints involved.

DR. JOHNSON: It depends on what you can get away with. If you can X-ray their hands at two hands, you probably should do it. I mean it would be interesting or maybe the other knee--

CHAIRPERSON PETRI: Well, I think if the structure claim is just before and after X-rays, I don't think we're increasing the cost of those trials very much to ask that there be--

DR. JOHNSON: I'm not talking about cost. I'm talking about the ethics of irradiating people.

CHAIRPERSON PETRI: Well, I'm not actually concerned about the ethics of a before and after set of knee, hip and hand X-rays.

DR. JOHNSON: Well, then you should run the trials. I mean that's the point. I mean I think these ethics change across investigators, too, and other people would argue against that, I think.

CHAIRPERSON PETRI: Let me ask the group. If for the structure claim, this is one, of course, we're going to commit patients to a drug long-term, don't we want information about multiple joints? Let me start with Dr. Madrid.

DR. FERNANDEZ-MADRID: Yes, definitely.

CHAIRPERSON PETRI: Dr. Callahan?

DR. CALLAHAN: It sounds reasonable. I would think you'd have to go through an IRB in terms of--

CHAIRPERSON PETRI: Well, that's a given. There is going to have to be full consent and approval of IRBs. Dr. Moreland?

DR. MORELAND: The answer is yes, but we're mainly looking at knee OA as a joint that we have good standard measures, "good" with quotes around it. If we go flipping up hand films and hip films, I don't think we have validated methods of looking at cartilage. So, yes, we want them, but we don't have the--

CHAIRPERSON PETRI: For hand films we do, don't we?

DR. MORELAND: For OA?

CHAIRPERSON PETRI: From the Baltimore Longitudinal Study of Aging. I thought there were validated measures.

DR. MORELAND: I'm not aware of the literature on that. So I don't know.

DR. LIANG: There are measures but not used in trials. I mean we have ways of saying what we see, but--

CHAIRPERSON PETRI: Well, but I think here we have the opportunity. I think that's going to be a terrible waste for this structure claim if we don't look at the other joints.

DR. JOHNSON: Well, intellectually I agree with you, sure, and these things have never been used in trials.

CHAIRPERSON PETRI: Dr. White, any comments?

DR. WHITE: I think the place where it is an issue, just as you say, with if you're going to go for a structural claim. Ideally I think it would be nice and we'd all like to know what you want to know, too, and the patients would like to know that. So if there is some way to incorporate it, it should be. Should be an absolute requirement? I'm not so sure given that you couldn't tell anybody how to measure it.

CHAIRPERSON PETRI: Dr. Harris?

DR. HARRIS: I think actually even with respect to a structure claim, the nonsignal joints--including that introduces, you know, another level of complexity. I mean structure itself, you know, we had a good degree of debate on it. Now, nonsignal joint involvement to me just introduces another level of complexity. That's my view.

CHAIRPERSON PETRI: But clinical trials are such an opportunity for new knowledge. So even if we don't have perfect validated measures for hips, I think we are quite close for hands. Why not take that, that opportunity? Because clinical trials right now are only probably going to be done in industry. I can't see an OA trial being done in an academic institution right now.

DR. HARRIS: Well, the point is that if we are looking at the knee and drug X and whether it responds structurally, you know, there may be ten patients who may only have a knee involved. You know there may be two who may have other joints involved. I just think that that puts another limitation on the drug, you know, that makes it more difficult to do. That's my view.

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: I agree ideally that it would be interesting information, but I think the issue of validation of these other X-rays and the fact that the standardization

of entry is going to be very different--obviously a knee arthritis protocol that people who enter are going to have a certain set of standardized parameters when they enter the study. And everything else is going to be very variable among them with regard to their other joints. So I think given the nonvalidation of the X-ray measurements in those other joints and the heterogeneity of the patient population, that it would be an almost impossible study to do.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: Nothing to say.

CHAIRPERSON PETRI: Dr. Pucino?

DR. PUCINO: It would be nice to look at the other joints but certainly wouldn't use it for a sample size calculations.

CHAIRPERSON PETRI: Comment?

DR. DOUGADOS: Just to summarize the discussion we had on this point in our group. The nonsignal joint, it was proposed we have two possibilities to evaluate the symptoms. The first one is to use a general tool. As an example, you say what about your other articular condition? And the criticism of this tool is that probably the weight of the spine was so important that there would be a big noise due to the spine. The second possibility is to use clinical

tools to evaluate all the different joints. That will be very complicated in order to conduct clinical trials. In other words, everybody agreed that, in fact, the structure to evaluate the other joints is the best thing to do.

I just want to remind you that from an ethical point of view, usually the ethical review board refused that you perform iterative pelvic X-rays in patients without hip osteoarthritis. In other words, if you conduct a study of knee osteoarthritis, usually you are not allowed to perform iterative hip X-rays. And that is the reason why. The conclusion was that if you're conducting a study of either knee or hip, at least you need to evaluate the contralateral, and if you want to have an idea on the effect on the general disease, keep also hand X-rays at baseline after two years in order to get an idea of the effect of the drug on the hand. But I assume it to be difficult to perform X-rays of all the joints at baseline and after one or two years.

CHAIRPERSON PETRI: Well, I can't imagine that we would want to recommend X-raying all the joints. I agree with you. Dr. Schwieterman.

DR. SCHWIETERMAN: This wouldn't solve everything, but there are a number of predecessors where you sub-study

within a larger study and you could conceivably consider something like that with this. It doesn't answer all the questions. There is obviously some concerns about that, but you wouldn't have to do that with everybody.

CHAIRPERSON PETRI: Other comments? The next question is number four: Should time be an explicit requirement for any claim or should any limitations of the data simply be reflected in labeling? I guess this gets to the question, for example, if the structure claim is based on one year, what do we tell the patient in the labeling? Dr. Madrid.

DR. FERNANDEZ-MADRID: Well, I think we should. My answer is yes, what we should tell the patient the results of the study, that this is not a promise of improvement in one year.

CHAIRPERSON PETRI: Dr. Callahan?

DR. CALLAHAN: I would think you should just list the limitations based on the time frame, list the time frame in the claim.

CHAIRPERSON PETRI: Dr. Moreland?

DR. MORELAND: I agree. I think the comments we've talked about before with pain and function being three months and structure being a year are fairly standard. A recent example of a drug that's been on the market that is a

pain medicine, anti-inflammatory drug that was only studied for a week or two weeks, I think now it's on the market, and questions of hepatitis from this particular drug. It wasn't clearly implied with samples that the patients get and so forth that this was only studied for seven days or 14 days. And I think high on the list of whatever that bottle is or the package is, that you need to say this drug has been studied only after so many time points. Because once a drug gets on the market and patients feel better, it's tough to stop it after three months, and then these unknown side effects. So I think not only do we need the label in the package insert and the label in the PDR, it needs to be very clear how long we as physicians should--and so patients can see that also.

CHAIRPERSON PETRI: Dr. White.

DR. WHITE: I agree.

CHAIRPERSON PETRI: Dr. Harris?

DR. HARRIS: Well, I think that there should be at least time, obviously, and I guess that's what time is about. So I guess certainly with respect to pain, certainly with respect to structure. While I couldn't--if we do separate pain and function, and I think we--I don't know if we came to consensus about that or not.

CHAIRPERSON PETRI: There's a question coming up.

DR. HARRIS: Okay. Well, you know, the one area in which I'm not, you know, very clear about in terms of putting a time is with respect to function. But I think there should be some, there should be some time written there, some least time.

CHAIRPERSON PETRI: Dr. Abramson?

DR. ABRAMSON: Right. I think the minimum amount of time is a very important issue to me. That as Dr. Moreland pointed out, if you get a drug on the market for an indication based on a period of time that might be inadequate to detect most if its potential side effects, there is a reason to be concerned, and I think in this instance the structure issue. We talked this morning about whether you could show differences in three months, would that would be adequate? And it might be adequate to show differences in structure, but I have concern given the nature of the kinds of drugs we're looking at that affect structure that you really need a year to look at adverse effects in terms of various substrates with these metalaprotenases, for example. So I think in answer to this question, I would encourage that we have a minimum amount of time for structure indication at a year even if differences can be shown earlier.

And then the other point which was discussed

earlier is what is a clinically significant improvement in structure which I think will await some of this analysis of outcome studies in OA.

CHAIRPERSON PETRI: I certainly agree with a structure claim requiring a one-year study, and I think it would be important actually to put it in labeling. But I wanted to ask you, Kent, has this ever been done before? Putting in a label that the claim is based on 12 months data, for example?

DR. JOHNSON: Well, Mike can correct me if I'm wrong here, but I think the trend is to more and more describe the actual trials. You know if you look at the cyclosporin label, even though the competitor drug is in there along with the placebos, you know, and I'm not sure that we can do much better than that. This issue of thinking about these damn things and trying to have a time scale that's adequate to pick up what is suspected is critical, but sometimes it still doesn't work.

CHAIRPERSON PETRI: It's going to be part of a class labeling though? What I'm asking now is if one study goes out to two years, can they state that?

DR. JOHNSON: Yeah, I think that should be--

CHAIRPERSON PETRI: We went to two years. This other drug went to one.

DR. JOHNSON: Yeah, I think so, and they could say they are a two-year structure improver as opposed to Company Y who only has a one-year structure improver.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: I don't want to be a wet blanket. I don't think time is the issue here. It's really--the first thing, you know, we're talking about something, some way to capture the effect and also some way to capture the bad stuff, and it was really how frequent it is and how many patients, you know, patient years have been exposed. It's really about, you know, if it's a drug that really hits the market like Phen Fen, you find out fast, and then if it's sort of not an interesting drug and no newspaper drug, you don't get very many people on it, and you never find out because it's a low rate of taking it. But there is no way, I guess, to say, you know, it should be a claim by patient years out or accrued. That's the more--that's what we're really interested in, I think, not the time.

DR. JOHNSON: Yeah. We have limited control over that. I mean the--

DR. LIANG: Yeah.

DR. JOHNSON: What are they called? The orphan drugs, you know, are one extreme.

CHAIRPERSON PETRI: Matt, I think I'm going to

disagree with you. I think we really cared about the bisphosphanate trials knowing when they had gotten to two years, three years, four years, you know, we felt so much more comfortable about keeping our patients on them long term.

DR. LIANG: I understand. I mean obviously I'm more comfortable the longer the drug is out, and I never prescribe new drugs. And I usually in my pitch to patients about drugs say that I have a lot of experience with this, a little experience or it just came out yesterday or it's been out 20 years. I mean I sort of put those in the qualitative part of my sort of informing patients in the office, but it's both. It's time and patient years in the denominator.

DR. JOHNSON: If your sentiment, Michelle, is that there is something about structure that is dramatic enough that it should have, you know, a longer time duration than just pain, let's say, I personally would agree with you, and that's kind of the issue. You know does the nature of the claim, should the nature of the claim mandate, you know, a certain time duration even if it can be shown in a shorter period of time. You know if MRI gets validated, that probably could show these things in three months or six months.

CHAIRPERSON PETRI: But again because we're

talking the slope this morning, since we don't know what's going to happen to that slope, is it going to plateau, is it going to--what's going to happen to it over time? I think for the structure claim especially I'm interested and I think my patients will be interested in how long patients have been followed to meet that claim. Other discussion? And then our next question--

DR. WITTER: Sorry, Michelle. Michelle, just before you move on, in terms of pain, would any of the committee's comments change given the consideration of the various modalities? I mean I'm just wondering if the comments haven't been all directed towards paras medications? I mean would any statements be different for topicals, for interarticulars? I just want to make sure that we've discussed.

CHAIRPERSON PETRI: Thoughts on that? Dr. Moreland?

DR. MORELAND: We haven't discussed interarticulars in this whole thing so I guess there obviously are some new devices and drugs that have recently been approved with interarticular in mind. And we're sort or stuck with what we have there and don't have some good ideas as to how move those along for how often to give them and so forth. So I don't think I have anything more to add

with regards to that unless we had more generalized discussion about those types of therapies.

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: Well, I guess the topicals in particular are a different kind of treatment in terms of the toxicity. One would think the toxicity issues would be substantially less and one would think that in the case of a nonsteroidal type drug, the reason one goes out to three months is more of interest in the potential side effect than whether you're getting a therapeutic response to pain. So I guess the topicals really holding them to a three month duration is probably too long, and the question is--there doesn't seem to me there's major toxicity to most of these drugs except for the local irritation. So the amount of time for pain treatment should be substantially less, I would think.

CHAIRPERSON PETRI: Any other comments? Dr. Witter, does that address your question?

DR. LIANG: I don't think that's fair though. You can't--I mean in the marketplace, in the office you're making choices between agents; right? So you can't--I mean it wouldn't be fair to assume that the three month topical isn't going to have some systemic effect. I mean what about steroids and cataracts now? I mean that kind of thing. I

mean it seems to me you should do it the same across the board because you have to choose between those.

CHAIRPERSON PETRI: Dr. Pucino?

DR. PUCINO: Yeah. I think it depends, again it's dependent on the type of, the class of agent that's used and bypassing the liver in the metabolism and those type of issues that would decide whether a topical should be for a short term or a long term.

CHAIRPERSON PETRI: So we're not quite sure whether that three month rule should always apply. The next question is I think the one that Dr. Harris wanted us to revisit from this morning, which is should pain improvement and function improvement be combined into one claim? And I'd like to ask Dr. Tilley just to repeat the comment she made this morning about separating pain and function domains.

DR. TILLEY: If I can remember it. I guess the comment that I made was that we were, we kept talking about the WOMAC combining the two, and I was saying that that was a measurement property and that if we had a way to talk about pain and function separately, that they could then be looked at separately, and basically it's a measurement issue, and I didn't see any reason why you shouldn't, but again that goes back to clinical. And I didn't hear any

clinical discussion to say that you shouldn't look at them separately. In fact, I heard clinical discussion saying you should if you had a way to define them.

CHAIRPERSON PETRI: Is there other discussion about this before I ask for people's opinions? Let me ask Dr. Callahan, do you feel that pain and function are important enough as separate domains that we should keep separate claims for them?

DR. CALLAHAN: Well, I feel like in particular, and most of my experience is in RA, but when you look at the correlations between pain and function, they range between about .6 and .8 so clearly they're highly correlated. But yet there are people who--and if you look at changes over time in pain measures and function measures, there are people who will improve in one and stay the same or not improve in the other. So I think a case can be made for they are distinct measures. I recognize that some of the measures like the WOMAC intertwine pain and function together. What I'm not clear is when it says--does this mean if the claim is made, they have to be confined or is it saying people will be allowed to make a claim that it improves either/or, or they have to improve both? What does this exact one mean?

CHAIRPERSON PETRI: Let me ask Dr. Johnson about

that. If you had a VAS for pain, and you won on that, but you lose on a WOMAC, for example, could you still get this pain and function claim?

DR. JOHNSON: Well, they were meant to be separable. If you win on the VAS for pain and you lose on the WOMAC subscale for pain, that shouldn't happen, I don't think. But if you win on the VAS for pain and you sort of trend on the WOMAC subscale for pain, but you dramatically lose on a function measure, I would argue that you shouldn't get the claim. The claim should succeed in what it claims to claim, but it should not deteriorate in the other one would be one way to work it. But what I heard this morning was that not just the measures don't confound these things but the actual concepts are confounded, and there is no way to unconfound them. You know you ask somebody about their pain by pain at sitting and pain at walking and pain at sleep. I mean at some sense at least some of those are functions.

CHAIRPERSON PETRI: Discussion? Because we obviously don't have a consensus on this so I need to people to tell us what they're thinking. Dr. White?

DR. WHITE: My view of this is that they should be kept separate because even though the measurements are not wonderful and we do know that they are very interrelated,

there may be instances and perhaps more in the future than you have seen in the past where they might be separable and they should be addressed that way.

CHAIRPERSON PETRI: Dr. Harris?

DR. HARRIS: My view is exactly the same, that really one should separate pain and function. And I guess to some of the discussion today, certainly we see a wide range of patients in clinical practice. I think the critical concern is relief of pain, but not necessarily, that need not be accompanied by improvement of function. But our patients feel better and we have achieved, you know, what we want to achieve. The facetious way of looking at it here is a chance to get two points instead of one, you know, one combined point, and then there may be some benefit in that.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: Ditto.

CHAIRPERSON PETRI: Does this actually mean we've reached a consensus on this, that we think that these are two separate domains? Is there any objection from the committee?

DR. TILLEY: I guess I'm getting the sense we didn't really answer Kent's question, which as he was restating it this afternoon, sounded to me like he was

asking if you got an improvement in one and a deterioration in the other, could you get an approval for the one that improved? Is that what you're really trying to get to?

CHAIRPERSON PETRI: That's a second issue. That's a linkage issue, isn't it?

DR. TILLEY: Yeah. Is that what you're really trying to get to, Kent?

DR. JOHNSON: Yeah, yeah. There is two questions. One is they should be separable at all? And the second one is if they are separable, should winning in one also co-require not deteriorating in the other?

DR. WHITE: If you require that, then they're not separable.

DR. JOHNSON: Okay. That's right. They're not totally separable.

CHAIRPERSON PETRI: I think the example we gave is if someone had total pain relief and then used the joint more, they might ultimately lose in function, and, in fact, it was an excellent pain drug.

DR. JOHNSON: Okay. But if they dramatically win in pain hypothetically and they dramatically deteriorate in function, we'll allow a little deterioration, let's say. That's why we were asking about trial size. Remember this morning? For the equivalence claim, that's implicit in

having co-success, you know.

DR. ABRAMSON: In that regard, Kent, I think it's important to try and understand why there is this discrepancy. I think to keep things clear, it's important to keep pain and function separate, and when you get a disparity between pain improvement and deterioration in function rather than no improvement, then you had a side effect potentially. Then you have to figure out why, why is that happening?

DR. JOHNSON: That's the other way of doing it. Don't require co-stabilization of the other parameter, but just look at it from a risk-benefit point of view.

DR. ABRAMSON: Right.

DR. JOHNSON: And maybe that's simpler in the end.

CHAIRPERSON PETRI: Dr. Strand had a comment.

DR. STRAND: Well, I have a question, and that is I don't understand how you can measure pain without measuring function or vice versa because, in fact, the WOMAC asks about functions in terms of pain. And it also asks about pain in terms of function. And, in fact, if you use the validated Lequesne or the WOMAC as a full score, you get pain and function. They are combined into the final score. You couldn't really deteriorate very far in pain and still win in the WOMAC unless your function had gone completely

off the map which would be highly unlikely. In other words, they're internally consistent.

The other thing is that as with say coronary artery disease, if your pain is decreased, then your function will almost always increase. It's a rare patient that wouldn't do that.

CHAIRPERSON PETRI: Well, but you can give the example that the pain decreased because the person had a myocardial infarction.

DR. STRAND: Well, that's fine, but that's an adverse event that does get reported and you do know about it, and we're saying the same thing here. It's almost like we're trying to separate these things. They may be separate domains, but they, in fact, are integrally related in how the patient would respond to the treatment, whatever the treatment is. And to try to separate it, I think, may be an artificial thing. A claim is a different thing that a domain. But in the context of how we see a patient improving from a therapy, we have to look across the domains that are important to them. And for some they may perceive it as pain and for others they may perceive it as function, but I would argue that, in fact, both are impacted.

CHAIRPERSON PETRI: I think what's happening is exactly what Dr. Tilley was saying is that all this

discussion when it revolves around WOMAC gets confused because this particular instrument and I assume the Lequesne as well have intertwined these things to the point where they may be difficult to separate out.

DR. STRAND: They have intertwined them, but I think everybody in this audience who has used these instruments tells will tell you that you can't separate them out. You really cannot. I mean a patient cannot, in fact, separate it out because if they choose not to want to walk across the room, then they may not have pain. And so it's quite specific to something we also call health-related quality of life, and it has to do with what their expectations are for their specific niche in life.

CHAIRPERSON PETRI: Dr. Singh.

DR. SINGH: I think I would tend to agree with you, Michelle, because we should keep in mind what we are talking about here. Are we talking about pain and function as separate domains or are we talking about the properties of a given instrument in being able to separate pain and function? And as you pointed out, it may be that that's how the WOMAC asks the question, but in order we've done things slightly differently. Dr. Callahan was just saying that the general correlation between pain and function, for example, instruments like the Health Assessment Questionnaire are of

the order of about .6 to .8. Yes, there is a correlation, but it's not 1.0. It's not perfect. In fact, we see it. We see it all the time that when we are looking at patients followed up for a period of time and measuring the pain with the VAS on the Health Assessment Questionnaire and the function of the disability index, the pain tends to improve a lot faster and a lot more than the functional one and function comes a little bit slower. Generally, yes, they go in the same direction. I have not yet seen a case, Kent, where you have a dramatic improvement in pain in any subgroup of patients and a dramatic decrease in function. I suppose it could occur. We haven't seen that.

But there are instruments that put the pain out of the domain, and they should be considered as separate domains, just like what you said.

CHAIRPERSON PETRI: Now, I gather that if this is done, this will be different from the European approach. Is that true? That it's combined in Europe?

DR. JOHNSON: Yeah. If we keep separate and they combine them, then it will be different.

[Laughter.]

DR. JOHNSON: I think the jury is hung on this one frankly.

CHAIRPERSON PETRI: Well, actually why don't we

take a vote just so you can get an idea about whether or not we reached a consensus. So the question is whether pain improvement and function improvement should be combined into one claim? Those who believe that it should be one claim, could you please raise your hands?

[No response.]

CHAIRPERSON PETRI: Those who believe there should be two separate claims, could you please raise your hands?

[Show of hands.]

CHAIRPERSON PETRI: So--

DR. SCHWIETERMAN: There you have it.

CHAIRPERSON PETRI: Now, our next question, number six, is it is best to leave how much clinical evidence of pain or function improvement is needed for a structure claim unspecified? So, Dr. Johnson, can you give us some background on this?

DR. JOHNSON: I think we've been through this. The more fundamental question is should there be any clinical evidence that you need if you've got something that--it sounds to me like actually the sentiment this morning was if we could have consensus on what's a clinically relevant X-ray change, that if a drug succeeded in doing that, the rest could be done in phase four, which would imply no clinical test at all at approval for a

structure agent.

CHAIRPERSON PETRI: So basically we're going to keep structure as a separate domain because this morning, I don't think we wanted to link these.

DR. JOHNSON: Yeah. I don't know if we can do it from a regulatory point of view, but we'll have to try.

CHAIRPERSON PETRI: My thought at least was that you could perhaps preserve structure, but what had already been lost would be the cause of pain and loss of function.

DR. JOHNSON: What do mean already been lost?

CHAIRPERSON PETRI: Well, prevent further loss.

DR. JOHNSON: Oh, okay.

CHAIRPERSON PETRI: A drug could meet the structure claim because it prevents further loss, but it might not do anything for the pain and loss in function that had already occurred.

DR. JOHNSON: Yeah. But the question is should there be any evidence, any clinical evidence, of the test arm being better or less worse than the control at the one-year time when you win by structure, and, you know, maybe we could argue that if there is a strong sentiment that there shouldn't also be a concomitant clinical test or the traditional clinical test, that it should just be a weaker test, you know. It should trend in the right

direction or something like that or it should not trend in the wrong direction.

CHAIRPERSON PETRI: Dr. Liang.

DR. LIANG: I didn't realize the implication of my vote. I thought that if we let you say whatever you wanted to say, you know, that you replaced the divots, you know, in the cartilage, that some day you would give us the goods on whether it made a difference to the patient.

DR. JOHNSON: Yeah, but the issue is are the clinical outcomes going to occur at approval or are they going to be deferred into a phase four study?

DR. LIANG: Well, didn't you just say that they don't have to give you that stuff in phase four, which I didn't realize?

DR. JOHNSON: Well, that's what we're debating.

DR. LIANG: Oh, oh. I'm saying, yes, you can, you know, if they can make a credible case and it stands on review, let--

DR. JOHNSON: In terms of what?

DR. LIANG: Well, if it's a rigorous study accepted, but then I think they would want, I would think, and you would want that they have a mandatory requirement to come back with the clinical stuff at some future point.

DR. JOHNSON: Yeah. It would be like the RA

accelerated approval business.

DR. LIANG: Yes.

DR. JOHNSON: Where there is something major on X-ray.

DR. LIANG: Yes, right.

DR. JOHNSON: And the clinical validation occurs phase four.

DR. LIANG: Okay. I'm comfortable with that at least for now.

CHAIRPERSON PETRI: Dr. Abramson.

DR. ABRAMSON: The other argument, the other way to look at it, though, is again just if you get approval for what you set out to prove and that information is valid and made public to physicians, they then would act upon that information, and if that information included not too much clinical benefit at one year, that would be part of their decision-making process. But I think unlinking these things is important to allow a drug to come to market based on the claim that it makes as long as it doesn't make claims that it didn't show to be valid. So I'm comfortable personally with a structure claim but with just some information what the clinical benefit was.

DR. JOHNSON: Well, it might be critical as to what you mean by "some." If it's not a standard hurdle,

maybe that's okay. You know I think traditionally the FDA has approved drugs because they improve signs and symptoms, not because they approve some imaging measure or some lab test. But things are changing.

DR. SCHWIETERMAN: I'm confused, Kent. It seems to me there are two questions on the table. One is it every necessary to have clinical data associated with the surrogate and the other one is when ought, if so, when ought it to be done? And I hear assumptions and implications that I don't quite understand.

DR. JOHNSON: Well, I don't know if anybody has argued that there should never be clinical data forthcoming. I think the issue is how much of it should be available at approval time, and if not then, you know, should there be a phase four scheme to capture that?

CHAIRPERSON PETRI: Other discussion? Dr. Moreland?

DR. MORELAND: No, I view what Dr. Abramson said, you know, the structural claim is fine, but then I'm going to make my decision whether to use it or not. I'd like to see some of the clinical data. If by chance the WOMAC improved greatly, then I'm going to be more inclined. If it wasn't quite so good, then I have to have that discussion with the patient, say we know that it changes the X-rays,

but it didn't change your symptoms. So I think you have to have that data, but I wouldn't tie that in to getting approval of the job. I think from my perspective if there is a good robust number that you choose for X-ray changes that we think is clinically important, then I would use the drug, but I'd like to have everything else to go with it to see if it did improve the structure, it even helped improve their function, too.

DR. JOHNSON: We'd probably even want to describe the clinical data, you know, the sort of provisional clinical data at approval time in the label even though it's not a clinical endpoint at that point in time.

DR. MORELAND: That's right.

DR. WITTER: Michelle, could I ask you to take maybe a quick vote of the members like you did before just to--

CHAIRPERSON PETRI: Certainly. Let me see if I can phrase a question here.

DR. WHITE: Could I just make a comment, Michelle, before we do that?

CHAIRPERSON PETRI: Yes.

DR. WHITE: We had a discussion at lunch, Barbara and I, and I still remain foggy about this one because it would seem to me that the sole reason that you would want a

drug that could preserve or a reason that you would want to have a drug to preserve structure would be because you wouldn't want it to impact upon what counted to the patient which was pain and function. I don't think the patients really care what happens to the structure if it doesn't do anything for their pain and their function. And I think that what we've been talking around is the fact that, yes, that's a separate issue, does it affect something structurally, but really what is going to count when we judge are we going to give that drug, are we going to recommend it to our patient, is whether because we affected structure, we had a beneficial effect on pain and on function.

So I think about this one perhaps a little differently in terms of does there then--maybe because it takes a long time, and it is reasonable to give an accelerated approval--but I wonder if, in fact, we shouldn't have in this case some requirement for follow-up that it makes the difference that we want and that's why we did it in the first place?

CHAIRPERSON PETRI: I think that's what we've been saying that in this particular case, we are very interested and think, in fact, it should be mandated that there be phase four data.

DR. JOHNSON: Yeah, an accelerated approval for structure would require, you know, under this standard rubric, it would require phase four clinical information, which if that information proved to be negative, then you'd presumably have to withdraw the drug.

CHAIRPERSON PETRI: Dr. Schwartz.

DR. SCHWARTZ: I don't really disagree with that. I think the only comment to be made is that we are really lacking any data to say when that increase or improvement in eventual pain and function would come about. You know we don't know if it will be at the time of approval. I think it's unlikely, but it's going to be five years or ten years or maybe 15 years down the road, and how do you know when you haven't achieved it? I mean there really are just no data on which to base when you'll be successful or not until you accumulate that over a long period of time. So I agree with the phase four studies, but I think it's going to be a pretty long phase four.

CHAIRPERSON PETRI: This is a tough issue because we want these drugs, and we don't want to delay their development and I think that's the reason that most of us feel very comfortable with having the structure claim at one year, but with phase four data mandatory. Dr. Madrid?

DR. FERNANDEZ-MADRID: It is a very tough

question. I think I like the structural claim. I think a structural claim is fine. However, in the absence of clinical data, I think I would be uncomfortable to approve the drug at that time. I think this is a big if for a surrogate like this because we know that many outcome measures give us data which is statistically significant but clinically not significant, and who is to tell me that these structural claims that have statistical significance have real clinical significance?

CHAIRPERSON PETRI: But are you comfortable with the structure claim followed by phase four data, and if the phase four data don't show clinical importance withdrawal of the drug?

DR. FERNANDEZ-MADRID: Well, I think in that case I would agree with you.

CHAIRPERSON PETRI: What I'm concerned about is if we require a five-year study to get the structure claim--I'm assuming five years would be enough to show clinical importance or maybe it would be ten years, I'm not sure we're going to see these drugs developed. We will make the hurdle too high. Let me ask Dr. Liang.

DR. LIANG: I'm sure everyone is going to collect the clinical data at one year. And no one, I would imagine, is foolish enough maybe to give it in asymptomatic

individuals. So I think we're just saying that at one year you can get away with just a structural thing and maybe poopy improvement in pain and function, but I think--and I'm sure they would want to do this. I mean you're not going to get anybody to take this unless you can provide that data so I don't think we're unleashing, you know, something out of the bottle that can't be put back in.

CHAIRPERSON PETRI: Well, this is an unusual way to phrase a claim, though; isn't it, Kent? The structure claim, one year, but then requiring phase four?

DR. JOHNSON: Yeah, but it's standard surrogate stuff, you know. I mean this is a better surrogate that CD-4 a priori in my mind.

DR. LIANG: I think one could get beaten up on what you present as the sort of morphologic data as to how likely it is to be something that would result in a distal benefit.

DR. JOHNSON: Well, what would happen is there would be a difficulty in continuing the trial because people would clamber for the drug and it would be poorly controlled, I'm sure, by two or three or four years out. I'm not sure that makes it uninterpretable.

DR. WHITE: Could I ask what is structure a surrogate for?

DR. JOHNSON: Eventual improved, eventual better or longer functioning, no joint replacement, you know.

DR. LIANG: Or stabilization of pain. It doesn't get worse.

DR. TILLEY: That's what I was concerned about because we kept talking about having to show some sort of functional improvement, but maybe it's enough to show that people just stay the same.

DR. JOHNSON: Yes.

DR. LIANG: Yes, I think that's fine.

DR. JOHNSON: We haven't talked about what the clinical test should be. We're just talking about clinical assessments. We haven't even specified what those are. I mean that would be the next question.

CHAIRPERSON PETRI: I think what we said is that the X-ray is not enough. It must have some clinical importance in the patient.

DR. WHITE: That's a good way.

CHAIRPERSON PETRI: Dr. Beary.

DR. BEARY: I think some of these points have been brought out here, but one of the practical issues if you're looking for phase four studies in the five or so year category that is going to throw a tremendous chill on looking at primary questions that deal with structure

because the expectation is that in early disease the pain and function, everybody will be looking at them in their trials. There will be data available at the completion of trials to address them, but it is entirely possible that those effects will be delayed. And these things get a bit tautological, but there are some excellent analgesics here. Right now there are no structural drugs whatsoever so I think the cautious way you're approaching this to make sure this particular field of development does not get chilled is a very useful thing for those of us who have to plan and justify these experiments back in the firm. Thank you.

CHAIRPERSON PETRI: Dr. Harris.

DR. HARRIS: Yeah. Maybe I'm saying the same thing again, but we've been able to deal with analgesia and pain relief in patients with OA. I think the challenge is that we haven't been able to prevent progression of the disease. Certainly if there are new agents coming along that are going to do that, then one could conceivably see in the brave new world that perhaps there may be two drugs, a disease modifying agent in which structure is important but that, in fact, taking a longer period of time to see its efficacy, and another drug that is analgesic. And so I feel that, you know, that as far as structure goes, certainly it needs to be decoupled and certainly in terms of a claim I

think it seems to be legitimate to use structure alone as a claim.

CHAIRPERSON PETRI: But, Dr. Harris, let me challenge you on this because if we let structure stand alone as a one-year claim, isn't it possible that we'll have a drug out there that, in fact, is not improving patient's clinically?

DR. HARRIS: Well, clinical in terms of--

CHAIRPERSON PETRI: Eventually doesn't it have to stabilize pain and stabilize function for us to want to use it? Dr. Abramson?

DR. ABRAMSON: I think one other way to get at that is that it may not be the studies in the structure change that show that. In other words, if studies of X-ray changes in OA show in other studies that changes of .2 millimeters a year are standard and the natural history of the disease will progress at a certain rate and you have a drug that prevents that amount of change, even without during the course of that one year showing clinical outcome, we may be in the position to approve a drug for structure and use parallel validated studies of what we know about the history of the disease to be able to make the leap of faith that if you stop that, that is going to have a good clinical outcome.

It may be that the patient population that we study, these earlier mild OA that might be very good candidates for this drug, won't have a lot of pain, and so it might be very difficult in that population to show pain relief. We might have to rely on other studies to allow us to make that leap of faith.

DR. WHITE: That's a good point because those are the patients that you would really like to treat before they ever get pain and loss of function. That's an excellent point.

CHAIRPERSON PETRI: Dr. Schwartz?

DR. SCHWARTZ: I was really going to make the same point that Steve just made. I guess I would disagree a little bit that the structure is a surrogate for pain or really any other kind of clinical marker, and I think when you look at an X-ray, you're really looking at a surrogate for cartilage destruction, and, in fact, if you are preventing the joint space from being additionally narrowed, you are preserving cartilage in theory. And again there is a leap of faith that if you preserve cartilage, you will preserve clinical function and hopefully put off pain as long as possible. So I don't really see why structure by itself could not really be a claim.

CHAIRPERSON PETRI: Dr. Schwieterman.

DR. SCHWIETERMAN: I think that there are some, I agree with some of that sentiment. However, I think that once you begin to think of scenarios whereby you would get minimal changes in structure, you would begin to question whether, in fact, that structural change that you had seen actually had any benefits to the patient. So I think it may be difficult in the abstract to discuss this because I think if you saw dramatic changes you could make that case. But more often than that, the case isn't dramatic and there are marginal differences and consequently a lot of unanswered questions, and I would think that physicians would want the clinical data in those cases.

CHAIRPERSON PETRI: We listened to some hypothetical examples this morning that perhaps you would be preserving the joint space with cartilage that wasn't good cartilage, for example. I mean we can think of some scenarios where you would still want that clinical correlate. Other suggestions or ways to reach a consensus on this? Everyone is thinking. The audience is thinking. Dr. Schwartz.

DR. SCHWARTZ: Well, I think what the conundrum that we're in is that we're trying to answer a question without really having any data at all, and we can be very hypothetical here and say poor cartilage or good cartilage.

Maybe poor cartilage is better than no cartilage. We really don't know. I think the thing that I would maybe just want to caution the committee about and I guess in the form of a plea for us in industry is not to make the hurdle so high that we don't end up trying to get these drugs to market. Because we're not going to be able to answer these questions until the drugs are out there and in clinical use and to see really what happens with them. So I think in the meantime a claim for structure would be a reasonable claim, and then we're just going to have to get the data thereafter to see if it really pays to do it.

CHAIRPERSON PETRI: Well, I agree with you. I don't think we want the structure claim to be longer than one year, that the issue is what sort of post-marketing phase four is going to be appropriate. I think it's Dr. Stephens?

DR. STEPHENS: Correct. Just as an additional item to clarify, it is indeed not an improvement necessarily in function but a stabilization or a reduction in the progression of loss of function or pain. Secondarily, one shouldn't think of just the number of millimeters in the joint space narrowing at one year. But since you are expecting to use these drugs over a long period of time, it's more like mortgaging of the joint space such that you

may have 30 percent preservation relative to placebo, but you have to look at that over time. And as time progresses, the relative difference between treated patients and untreated patients would be expected to grow, and so even though a small change at one year may not be very impressive, if you look at that change over the duration of the disease course, then that actually comes out to be substantial.

CHAIRPERSON PETRI: Other comments? Dr. Madrid?

DR. FERNANDEZ-MADRID: I think I was convinced by the arguments of Dr. Abramson, but I would be very surprised if we have a good drug that produces structural changes in one year we will not see a clinical counterpart of this. I suggest that this will happen.

CHAIRPERSON PETRI: Positive thinking. I'm going to phrase this vote with these two choices. One is that the structure claim at one year stands alone. And the other is that the structure claim at one year is coupled with the clinical correlation such as stabilization of pain and function done as part of phase four studies.

DR. ABRAMSON: Can I ask a question?

CHAIRPERSON PETRI: Yes.

DR. ABRAMSON: When you say phase four studies versus just clinical follow-up, what are the implications in

terms of costs and doability of that?

CHAIRPERSON PETRI: Well, let me ask Dr. Johnson.

DR. JOHNSON: Well, I think the implication is that a formal phase four study implies a control so that's an extension of your ongoing studies or a whole new study. So I'm sure it's substantial because I'm not sure what the interpretability of open data would be given this sort of non-robust character of the epidemiology we have right now.

CHAIRPERSON PETRI: I think one thing we're not doing is saying how long that phase four would have to be.

DR. JOHNSON: That's right. Yes, Ben's point is very well taken.

CHAIRPERSON PETRI: So if at 18 months that clinical correlation becomes obvious--

DR. JOHNSON: I know you can use huge trials to try to shrink the time. Or if you really have an incredibly slow-acting drug, maybe it will take three or four or five years in which case that's probably totally infeasible. I've not got that impression that the times are so long in talks with the companies, however. And, of course, you sequentially enroll so you've already got a bolus of patients who have been out maybe two years by the time all of the patients get out one year. So you've got, you're sort of halfway there.

CHAIRPERSON PETRI: Dr. Moreland?

DR. MORELAND: Why did you put the stabilization in the phase four? Why didn't you put that in the context of the clinical trial that got the structural claim?

CHAIRPERSON PETRI: I think it's very possible that it does not have to be done in phase four. If someone has the structure claim at one year and they already have the clinical correlation, that's fine. They don't have to extend into phase four. So maybe I should say the stabilization of pain and function shown during the trial or in phase four--a better way to restate it. Other suggestions on restating the question or other discussion?

DR. WHITE: I think that there may be problems with use of stabilization because if, in fact, ideally you would use this in patients to prevent ideally before they have much pain and function, then stabilization wouldn't be--I mean zero is zero or, you know, a little is a little that you would want to prevent. So I don't know that stabilization, it's a component of it, but maybe it's not the sole component of what you--

DR. JOHNSON: It's a clinical separation from your control. I mean both arms might deteriorate, but one deteriorates less fast than the other.

DR. WHITE: Were the patients getting the drug

doing better?

DR. JOHNSON: Yeah, yeah, yeah. Compared to control. That's always implicit.

DR. LIANG: Can I throw out something that I think has been seen in the chondrocyte business is that the structural advantage you got at one year is gonzo later on, and your pain is better. What do you do with that?

DR. WITTER: Can you repeat that?

DR. LIANG: I mean a lot of things have, you know, trajectories and it's possible that the effects you have on the cartilage are short-lived or it actually makes the stuff worse after time, after initial improvement, and you would get--

DR. JOHNSON: So you're saying one year might not be long enough?

DR. LIANG: No, no. And then it's conceivable because we think there is such mismatch between structure and symptoms that someone is better symptomatically and yet his structural indices are worse.

DR. JOHNSON: Well, I think it's always possible that the one year call for structure is the wrong call, and maybe it should be a two or a three year. In fact, there is something about osteoporosis that mandates a longer call. I can't remember what the explanation for that was. But,

given the sentiments that we've heard today, I doubt if making it a longer call is going to carry much weight. So that possibility I think is inherent in the process and there is no solution to it.

DR. LIANG: Well, would you approve it?

DR. JOHNSON: Well, we probably would approve it. And the market would eventually realize that it's worthless after a year and a half or two years and stop using it presumably.

CHAIRPERSON PETRI: It's hard to make those decisions through the market though. It's much better to do them scientifically.

DR. JOHNSON: That's true, but I don't--but you could make that argument about any drug studied for any period of time really.

CHAIRPERSON PETRI: Dr. White.

DR. WHITE: Just in terms of wording, Michelle, perhaps I would be more comfortable with something that you would get a structural claim, it's separate, it can be done at one year, but that there also has to be some evidence of clinical benefit, be it pain and/or function either during the trial or after the trial.

CHAIRPERSON PETRI: I think that's fine. Dr. Harris.

DR. HARRIS: Can I make--it seems to me that even, let us suppose that the agent improves structure, but it does so in a way in which it takes three months, six months, in fact, before there is indeed any significant improvement. What do we expect to happen in trials with the patients three to six months out? Do they stop all drugs? Do they go on nonsteroidals during that time or rather analgesics in which case making an estimate of this modifying agent and its effect on pain and function becomes difficult? So it seems to me that linking it certainly in the short-term, you know, structure with pain and functional improvement, if this indeed has a slow response in terms of its effect on structure and even, you know, affecting structure may, indeed if structure is what drives pain, then one might expect indeed that the effect on pain and function might be delayed. What happens?

Do you keep your patient then for six months on this agent alone and, you know, what do you do about pain and function during that period of time? It makes it more difficult is what I'm trying to say.

DR. LIANG: Most likely all patients are going to be on background therapy, you know, maximal background therapy or some sort of constant background therapy, and they're still going to have dysfunction and pain as a

consequence, so you got some variables that will move, and presumably your drug when it eventually kicks in will move one arm and not the other arm in your endpoints. And background therapy is sort of--I mean it will get confounded by that if you allow adlib use, and you're going to have to account for it, and you may have to allow adlib use from an ethical point of view. But I think that trials are doable if you get the right patients in there. You may not be able to get real mild ones in there because Tylenol will obliterate their pain. It would be a bad candidate for the trial.

CHAIRPERSON PETRI: Dr. Egger.

DR. EGGER: I think it's important to expedite research in this area, but I want to express some concern about voting for an alternative that includes a stage four study because I'm not sure that a stage four study could be definitive, and I think if I were sure of that, I would be much more comfortable. The particular thing that is bothering me is we've seen tremendous placebo response in the cooperating clinics in osteoarthritis and in, well, in rheumatoid arthritis I guess I'd have to say.

And if you have placebo responders who are, they're survivors. They're the people who could stay with an ineffective treatment that long. You may not see the

relative difference relative to placebo growing in a study group. These placebo responders may be getting better and better. I feel like it's very--we're talking about these stage four studies and they're kind of vague. I don't see a design firmly in my head. I'm not sure that there would be a problem, but I'm very uncomfortable voting for something that includes a stage four study when I can't see for certain that it would be definitive.

CHAIRPERSON PETRI: I can't think of another study design, though, that's going to allow a definitive answer. Let me ask Dr. Johnson if he has any other thoughts.

DR. JOHNSON: Well, you can double blind withdraw from any study actually if you want to sort of reaffirm what you thought you saw was really true, but I would have fancied that what's going to usually happen is the pivotal trial or trials are just going to be continued into phase four. So the whole design will already be done, and you'll be halfway there, and you're right, I mean if you've got some differential dropout problem, it could confound the result especially if the result is a small one to begin with. It could undermine the result, but I don't think that that's any different than those analytic challenges in any other setting. Why would it be different here other than everybody wants to get on the wonder drug? But at least for

the first one out of the shoot, there is no wonder drug other than the drug that they're already on, and since it hasn't affected symptoms, and you're really using symptoms as an endpoint, hopefully you can keep enough of them to the end of the pike.

It may be after one of these is approved that ethically that drug will have to be part of background therapy or that they'll have to do an active, I mean a similarity design rather than a difference design. But it sounds like for the first one out of the shoot, you would just continue your pivotal trials.

DR. EGGER: I'm wondering if you approve a drug based on a structural claim and you do a stage four study and there isn't, you can't show definitive clinical improvement or non-deterioration or whatever, and there are methodological issues and people generally believe that it would have been very hard to solve those methodological issues, would you then withdraw approval?

DR. JOHNSON: But why are there more methodological challenges just because the trial design happens to lapse over the approval time? I mean you're right. I mean a failed trial could always be due to a failed design rather than a failed drug. But it doesn't strike me that this is any different than any other

scenario; is it?

DR. EGGER: I think qualitatively the issues are the same. In terms of duration of the study, the longer you have a study, the more likely they are to occur.

DR. JOHNSON: Yeah. I'm nervous about five year studies.

DR. EGGER: Yeah.

DR. JOHNSON: I'd rather triple the sample size and make it a one-year study.

DR. SCHWIETERMAN: Let me just address your concern. I think that there is validity to your concerns because traditionally phase four studies have been less rigorously performed than the premarketing studies because there is less incentive oftentimes to do this. Under accelerated approval, there has not been a whole lot of experience with it, but presumably there would be more attention to the rigor given the possibility of having your product withdrawn. But I just wanted to affirm that your suspicions, in fact, are some of my suspicions as well, given my experience.

CHAIRPERSON PETRI: Dr. Moreland first and then Dr. White.

DR. MORELAND: I still think you're going to have

major difficulty in a phase four study of not giving everyone the real drug. So if you roll your pivotal study into phase four, you've already lost your control group. So do we come back and not get so excited about MMPs and say let's do the one-year study and if it's statistically significant go back and repeat it. Because I think it's going to look bad on any agency or anyone if we put a drug out and then say, oops, we didn't do the right study. We have to take it off the market now.

DR. JOHNSON: Well, yeah, but that's, you know, that's possible right now as a matter of fact. Most of the withdrawals have been for toxicity. I guess maybe all of them have been, but, you know, if you start a whole new study, you're going to have the same issue. You can't use a negative control; right? Everybody is going to want the drug.

DR. MORELAND: But once you have it on the market, it's tough to withhold the real thing, but if it's still not on the market you can put them in the placebo control trial because you could argue that you haven't shown in a scientific rigorous manner that it is effective. You haven't done the two studies.

CHAIRPERSON PETRI: You're talking about a two-phase approval; is that what you mean?

DR. MORELAND: I'm just, as the discussion goes on here, I'm having a little more trouble with if we're going to tie in that we need some improvement in function or stabilization in function and signs and symptoms, the way to really prove that is not in a phase four study where you roll patients over during that one pivotal study. You're going to have so many confounders. I think it boils down to are we going to trust the one structural study to be enough to get it on the market? And I think is it going to be one or do we need two? That's sort of where I'm coming from as a gestalt. I know we're all excited about moving this field along, but I'm equally not excited about telling patients that we didn't have design studies and we had to withdraw the drug because we found some bad side effect or we repeated the study and extended it and it didn't work.

DR. JOHNSON: So but you're comfortable approving something with its clinical importance never validated in a formal way?

DR. MORELAND: I didn't say that yet.

DR. JOHNSON: Didn't you?

DR. MORELAND: No.

DR. JOHNSON: Okay. I thought that's what you were saying.

DR. WHITE: Well, that's what I was going to

suggest. I mean it actually is the opposite end of what I did after hearing all this discussion. The question is to the rheumatologists around, if you had a drug that you knew gave what you believe to be reasonable salvation of joint space, would you give it to your patients in the absence, after you discuss costs, what was known about side-effects? Would that be enough for you to feel comfortable that you think you could make that leap that it would then be good for you patients? Is that enough goodness? Is that enough value to just let that be it for giving it to a patient?

CHAIRPERSON PETRI: And, for example, would you give that drug to your 30-year old who had a very strong family history of osteoarthritis? And I think the answer is probably no without some evidence of clinical benefit.

DR. JOHNSON: It may be that the evidence will accrue in five or ten years and we get more epidemiology, you know, and the decision would have been seen as wise ten years hence, but you would probably have to give it for 20 years, too.

CHAIRPERSON PETRI: Dr. Madrid.

DR. FERNANDEZ-MADRID: I think I would differ with you. I think the answer would be probably yes. It would be used. I think people would use it initially, and I think the same has happened with the drugs for rheumatoid

arthritis. In rheumatoid arthritis, there are excellent drugs, medium, medium effective drugs, poor drugs, and what happened is that people use those that are felt that are most effective. And those weak drugs are seldom or never used, and I--

CHAIRPERSON PETRI: But in rheumatoid arthritis, you get a more rapid feedback; don't you? The problem with these drugs is we're not going to be taking care of that patient in ten years.

DR. FERNANDEZ-MADRID: In rheumatoid arthritis, it was based on really lowering the standards, the AC-20, for instance, you go from a joint count from 20 to 16, for a morning stiffness from two hours from one and a half hour. The patients are still significantly, are significantly active in spite of the drug. And I think with this drug, if it's approved with this claim and nothing happens, it will die in no time.

DR. LIANG: Well, no one is going to buy it, no one is going to prescribe, and no one is going to buy it if it doesn't have that kind of clinical punch. So I think what we need is an incentive for the companies to hang in there. Could we extend the patent period? I mean I would because I think this is an important thing, and we have, you know, but I think that they realize that no one is going to

buy this or use it if they don't get that data.

CHAIRPERSON PETRI: I think Dr. Abramson was next.

DR. ABRAMSON: I just want in response to Barbara's question, I would prescribe that drug with a qualification, being that the amount of change that was demonstrated. In other words, I would accept the structure indication alone so long as the magnitude of the change in some way I became convinced was significant. And that's why I get back to the other issue. Someone is going to have to begin to tell us, based on the natural history of this disease, what begins to become significant changes in the cartilage. The only other comment, I would ask Larry in the tetracycline study right now where you're looking at the contralateral knee, that is largely asymptomatic--

DR. MORELAND: Yes.

DR. ABRAMSON: In the tetracycline study, the outcome there is not going to be an improvement of symptoms necessarily. It might be the prevention of osteoarthritis of that knee. Isn't that a structural indication?

DR. MORELAND: That's correct. This is to look at a structural indication and not a symptom. We're looking with symptoms also, but that's not the primary.

DR. ABRAMSON: But you would be satisfied with a structural outcome or the absence of development of OA?

DR. MORELAND: Yes.

DR. JOHNSON: Yes. We sort of separated that conceptually and that's why we put that in a totally different claim.

DR. ABRAMSON: Isn't it similar?

DR. JOHNSON: It is similar, but this is more preventing new disease. If you could do a study, you do it --strikes me as more, as a very persuasive structural endpoint as opposed to reducing your joint space narrowing by .01.

DR. ABRAMSON: It's probably not new disease. You would know better than me. But if you arthroscope that contralateral knee, it probably has osteoarthritis.

DR. JOHNSON: New clinical disease. New symptomatic disease, you know.

DR. WITTER: Could I ask for feelings on clinical personal experiences or trials that address if a compound were to arrest joint damage, the damage that's there? And I know you've discussed it to some extent. How that might persist even if the disease were arrested at that point in time? What one might expect for symptoms in terms of pain with that joint or any other nonsignal joints in personal experiences from trials and such?

CHAIRPERSON PETRI: Dr. Moreland.

DR. MORELAND: I don't think we have any good experience to answer that.

CHAIRPERSON PETRI: Dr. Liang?

DR. LIANG: I have none.

CHAIRPERSON PETRI: I'm going to go ahead and try to phrase this question. I realize we have not reached consensus. One possibility is that we accept the structure claim alone. A second possibility is that we couple it with evidence of clinical benefit either obtained within the trial or as part of phase four. Those of you who believe that the structure claim should stand alone, could you please raise your hands?

[Show of hands.]

CHAIRPERSON PETRI: Those of you who believe that the structure claim should be coupled with clinical benefit either as part of the study or as part of phase four, would you please raise your hands?

[Show of hands.]

CHAIRPERSON PETRI: So there is a split. But at least you've heard the reasons for the split, and I think it's very important for me to repeat that the committee is unanimous in feeling that we need drugs developed in this area and we don't want to hold back drug development. So I think that is the overriding sentiment that you heard, but

those of us who are clinicians also have some concerns about prescribing these drugs long term without evidence of clinical benefit.

Now there was one last question, and I think we may be able to deal with it very quickly because we discussed it somewhat this morning. Do you see insurmountable obstacles which in principle will make designs for claims of delay in new OA development and delay in surgical joint replacement fatally flawed? I think the second part we can almost dismiss. In our country and so many other countries, surgical joint replacement is not necessarily based on the severity of the OA, severity of pain and function, but what about claims of delay in new OA development? Are there thoughts about how those studies could be designed? Comment?

DR. DOUGADOS: I just want to come back to the problem of the delay of surgical joint replacement and just to comment on the joke of Matt this morning concerning the income and the outcome. Personally I am quite sure that it is a good outcome, and I will explain why. Usually when you are conducting the transsectional study looking at the reason for the indication of knee or hip replacement, you find a list of some reasons which have nothing to do with medical reasons, just for the surgeon, for the country where

you are working. But if you are looking at the epidemiological longitudinal studies in which a baseline you have enrolled patients with hip osteoarthritis, if you take the information concerning the structural severity, the symptomatic severity, and then you conduct longitudinal follow-up study. And I have in mind two studies, one conducted in UK by Michael Dougherty, which has been published in the Annals of Rheumatic Disease, and another one we are conducting in France. What we found, if we are looking at the probability of hip replacement and if we are conducting a study in which we are looking at the predictive factors of hip replacement, what we found is that symptomatic severity at entry plus structural severity at entry. Moreover, in the three year longitudinal study we have conducted, if we split the study in two parts, the first one of one year duration and the second part with two subsequent years duration, what we found, and that is an answer for the X-ray of the surrogate marker, and I come back to what I have said this morning, if there is a change, a structural change, within one year, that is highly predictive of hip replacement during the two subsequent years. In other words, I think that from data published in the literature, I think that we cannot forget this claim such as in proposal or at least this information seems to be

of clinical importance not only for the income of the surgeon but also for the quality of life of the patients.

CHAIRPERSON PETRI: I think your points are well taken, but it's going to be too difficult, I think, to make this into a claim. For example, as a clinician, I can tell you I have several patients who I think would benefit immensely from knee joint replacement, but they refuse because they're afraid of the surgery. I don't think this is an outcome that can be an objective one.

DR. DOUGADOS: No, I'm not--because I think there is a misunderstanding between claims and domain to be investigated. I agree that perhaps we don't need a specific claim, but at least that is a domain to be investigated to put in the dossier because I can tell you that in the field of, in the British study, 50 percent of the patients underwent hip arthroplasty after two years because it was [?]. In our study, 23 percent of the patients after three years underwent hip arthroplasty. But it is a huge amount of hip arthroplasty in the short period of time, only two or three years. So the description I have is not to propose a specific claim but at least a specific domain to be investigated.

DR. JOHNSON: Are you saying it's a scenario that's susceptible, that's possibly susceptible to trial

design? Randomized trials? Because I think the objections this morning were that there were so many non-medical confounders that you just have too much noise?

DR. DOUGADOS: Again, I'm not sure that it's so non-medical confounders. There is, but if you are looking at the VAS for pain, if you are looking what does it mean the absolute value of VAS, a lot of people think that it's not a pain VAS. It's a VAS related to a lot of things such as the particular sociological statutes are in financing, also the VAS. But no, I don't think it's possible right now to propose a randomized clinical trial with the primary criteria based on hip replacement, but I think that that will be at least of late clinical relevance.

CHAIRPERSON PETRI: Kent, I'd like to address the first part of your question, claim for delay in new OA development. I think the problem right now is there are no surrogate markers. A study for the delay in new OA development would have to be so long. So I think what's going to hold this up is going to be lack of surrogate markers.

DR. JOHNSON: Yeah. Well, I guess like Bill said this morning if a company disagrees and they come in with a design and they've shown it, we're not going to turn it away. There are a lot of issues that were brought up this

morning about surveying various joints and so on, but I mean I think this is what Ken Brandt is doing essentially.

DR. MORELAND: Yes, that's right. His study will I think be a test to see whether that's a doable study.

DR. JOHNSON: We'll learn from it.

CHAIRPERSON PETRI: Can you describe for those of us who don't know what his study is?

DR. JOHNSON: Well, Larry might be able to do it better. I haven't looked at it in awhile, but he did send it to me awhile ago.

DR. MORELAND: It's women who are between the ages of 45 and 60 who are anywhere from mildly to a lot overweight who have symptoms in one knee and have mild radiographic changes of one, no more than covering grade one, who are asymptomatic in the other knee and have essentially no changes, and are randomized, receive doxycycline and a placebo, and the outcome is going to be in the one that doesn't hurt at the present time to see whether that progresses to OA, measured by specialized X-ray films. We're getting bone scans to look at any possible inflammatory components that may predict.

CHAIRPERSON PETRI: Do you know what the time course of that study is? Five years?

DR. MORELAND: Well, if we were on course for

enrollment, it would be five years. It's going to be a little bit longer than that, I think.

DR. JOHNSON: But the duration of treatment is not five years?

DR. MORELAND: The duration of treatment is two years.

DR. JOHNSON: It's two years, yeah. I mean it's true I think there is only one or two cohorts of patients to use to drive the power calculations for these trials.

DR. MORELAND: Dr. Spector's trial with 40 some patients was the hypothesis behind this particular design.

CHAIRPERSON PETRI: Other thoughts about study designs to show delay in new OA development? Okay. Now, Kathleen Reedy is going to make an announcement about our next meeting.

MS. REEDY: The next meeting of the Arthritis Advisory Committee will be March 24 and 25 with all of the consultants also. And it will be at the Gaithersburg Holiday Inn, Two Montgomery Village Avenue, in Gaithersburg. On March 24, the committee will discuss its general scientific discussion, safety issues, gastrointestinal tolerability, renal, bone and reproductive toxicity, related to nonsteroidal anti-inflammatory drugs, for example, Cox 2 and other agents, with some representation from the

Gastrointestinal Drugs Advisory Committee.

On March 25, the committee will discuss the pain claim structure for chronic and acute pain and onset, fast onset of pain relief, including appropriate study design for prescription and non-prescription oral analgesics with guest experts from the pain expertise community and our consultants and representation from the Non-Prescription Drugs Advisory Committee.

CHAIRPERSON PETRI: I'd like to ask both Drs. Weintraub and Johnson whether they wanted to make any closing remarks in summary?

DR. WEINTRAUB: Yeah. I would like to make a closing remark. Actually the next, perhaps the second day of the next meeting will be much like today so I was thinking we could dim the lights and serve alcoholic beverages, but in fact the discussion this afternoon in particular was very lively and very enlightening. And I want to thank everybody for joining in.

CHAIRPERSON PETRI: Dr. Johnson.

DR. JOHNSON: Yes. I won't speak for the other FDA members, but I personally found this incredibly useful and I'm appreciative for everybody who came and participated.

CHAIRPERSON PETRI: And as always I'd like to

thank the committee members. The meeting is adjourned.

[Whereupon, at 4:15 p.m., the meeting was adjourned.]

- - -